コンピュータビジョンの今を映す
ーCVPR2017速報よりー

片岡裕雄

原健翔, 阿部香織, Yue Qiu

鈴木亮太，鈴木智之，大喜周平，張雨辰，範龍男

https://sites.google.com/site/cvpaperchallenge/

コンピュータビジョンの今を映す挑戦

論文読破・まとめ・アイディア考案・議論・実装・論文執筆・（社会実装）に至るまで広く取り組む

- 人員：産総研，筑波大，東京電機大，慶應大による約20名
- 読む：多読・体系化・共有
- 書く：洗練された問題設定を考案・論文化（・社会実装）
- 2017年はトップ会議※への投稿を目標に活動中

※ HP, Twitter, SlideShareもご覧ください

HP: https://sites.google.com/site/cvpaperchallenge/

Twitter: @CVpaperChalleng

SlideShare: @cvpaperchallenge

※ Google Scholar Top-20にリストアップされている国際会議や論文誌

@2015

CVPR2015の論文

計

602

本を完全読破

【主な業績】

・招待講演 x 1（PRMU研究会）

@2016

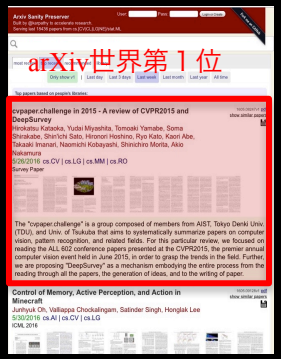
論文読破

計

1,000

本を達成

【主な業績】



- ECCV2016 WS x 2
- BMVC2016
- ViEW2016 oral/spotlight oral
- 招待講演 x 2（東大, IAIP）

@2017

TOP会議（GスカラーTOP20論文）への投稿

計

10

本を達成

【主な業績】

- ICCV 2017 WS x 1
- MIRU 2017 (oral x2, poster x1)
- MVA2017 x 2
- 招待講演 x 4（中京大, 慶大, 早稲田大, ABEJA）

ではなぜ、そこまでサーベイするのか？

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！

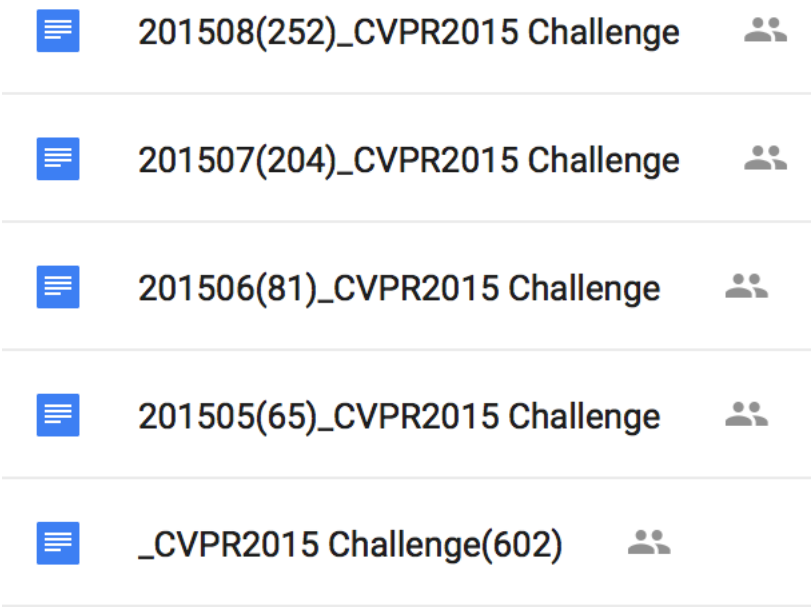
- CHIはユーザインタフェース界のトップ会議
- 2006年から12年間続く取り組み
- 日本のHCI分野の好調を支えている（CHI2017は日本がアジアで1位）

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．

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．

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．

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

CVPR 2017 参加報告

ほとんどCVPR 2017 速報の抜粋ですが，講演者が重要と思うものをピックアップしました
「CVPRとはどんな学会か？」よりも転じて現在のCV分野の動向を中心的に紹介しています

CVPRについて

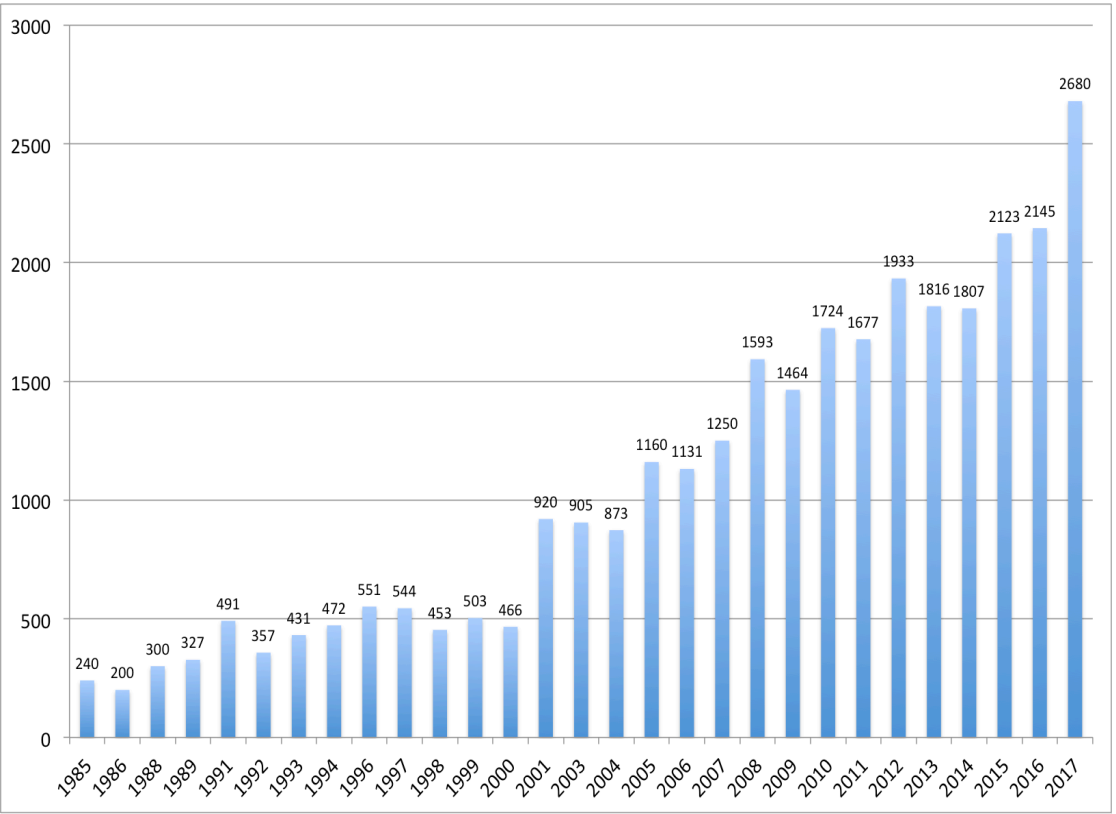
コンピュータビジョン，パターン認識分野のトップ会議

- Google Scholar (CV&PR)の統計にて常に一位
- 採択率は25~30%
- 2017年はハワイ・ホノルルにて開催

CVPRについて

AIブームの流れで投稿数・参加者数等増加傾向

- 投稿数は過去最高（2,620論文）を記録
- 参加者数 約5,000人，スポンサー数127



http://img.mp.itc.cn/upload/20170727/a94c7cfda6c34d2fl
c6b928_th.jpg

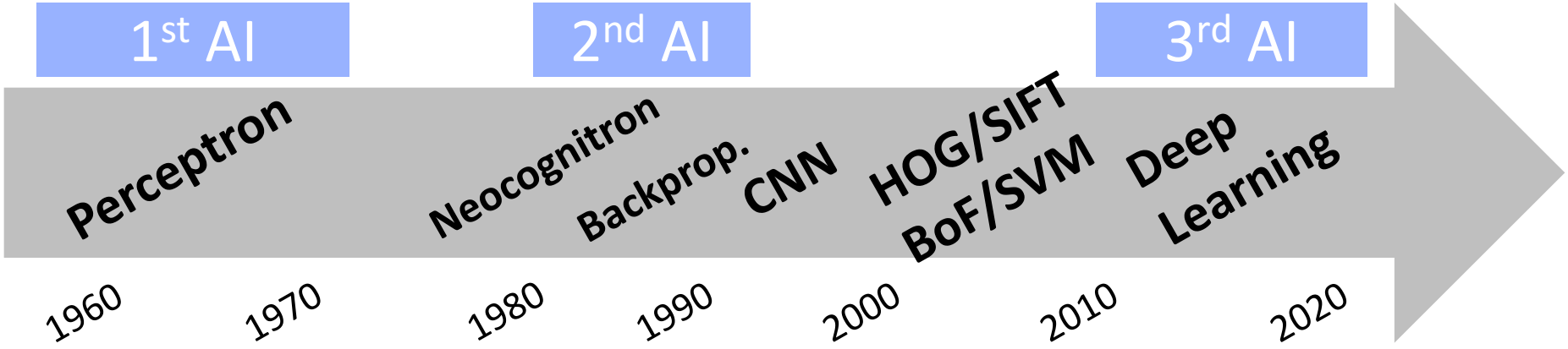
その前に（かなり大雑把な）DNNの概要

特にCVPR 2015あたりからあらゆる技術がDeep Neural Networks (DNN)により置き換えられた
CVPR 2017においては8~9割くらいの論文が同技術を用いている

DNNの動向（1/8）

DNN時代以前の動向

- Perceptron, MLP, Neocognitron, BackProp, CNN
- DNNが流行る前の画像認識では局所特徴が使用



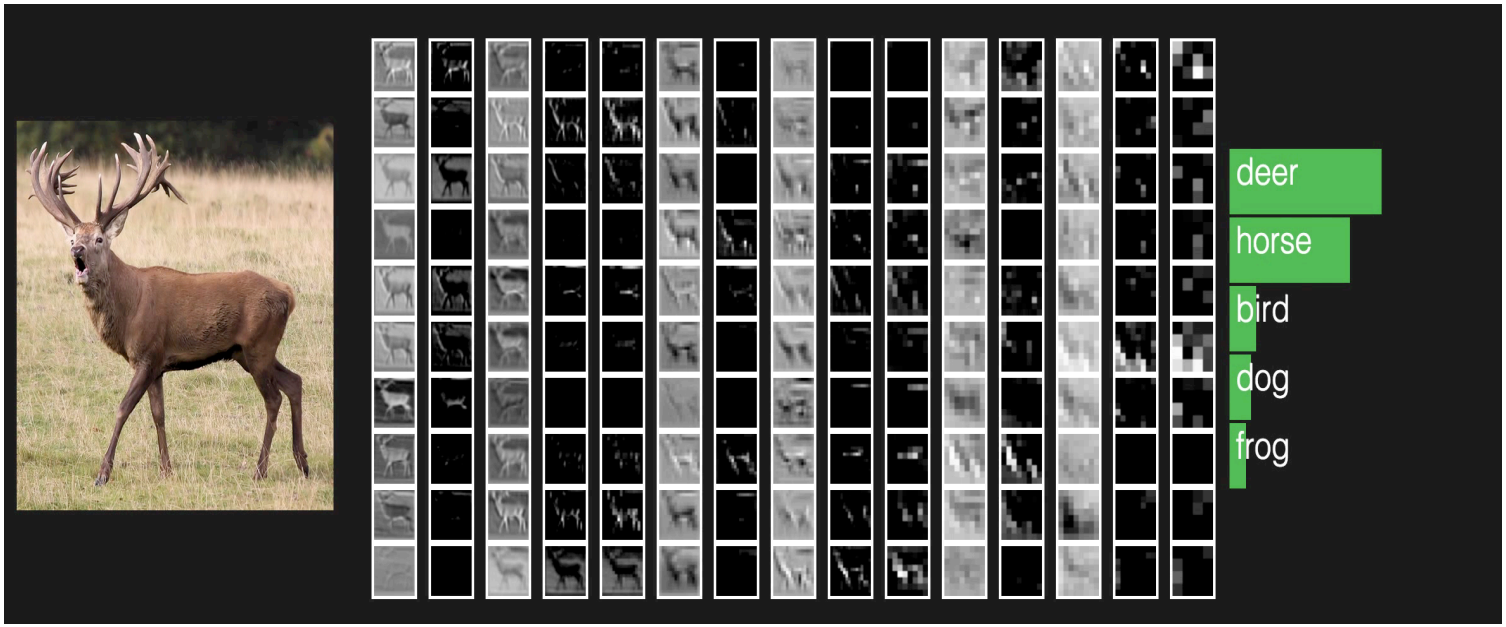
F. Rosenblatt et al. “Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms” in 1961.
Rumelhart et al. “Learning representations by back-propagating errors” in Nature 1986.
K. Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”, in 1980
Y. LeCun et al. “Gradient-based learning applied to document recognition” in IEEE 1998.

DNNの動向（2/8）

ILSVRCを発端とする画像識別タスクへの応用

– AlexNet @画像認識コンペILSVRC2012
  - 第一著者Alexさんのネットワーク（仕掛け人のHintonNetになってたかも？）

– 背景にはBelief Propagation, ReLU, SGD, Dropoutなど構造をDEEPにする技術が揃ってきた

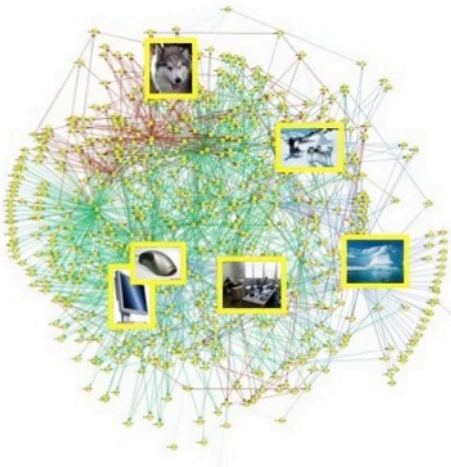


https://www.youtube.com/watch?v=zj_JlVqWK1M

DNNの動向（3/8）

DNNが勝てた背景

- ImageNet!（データが最も重要）
- NVIDIA!（圧倒的な計算力）



IMAGENET

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

×

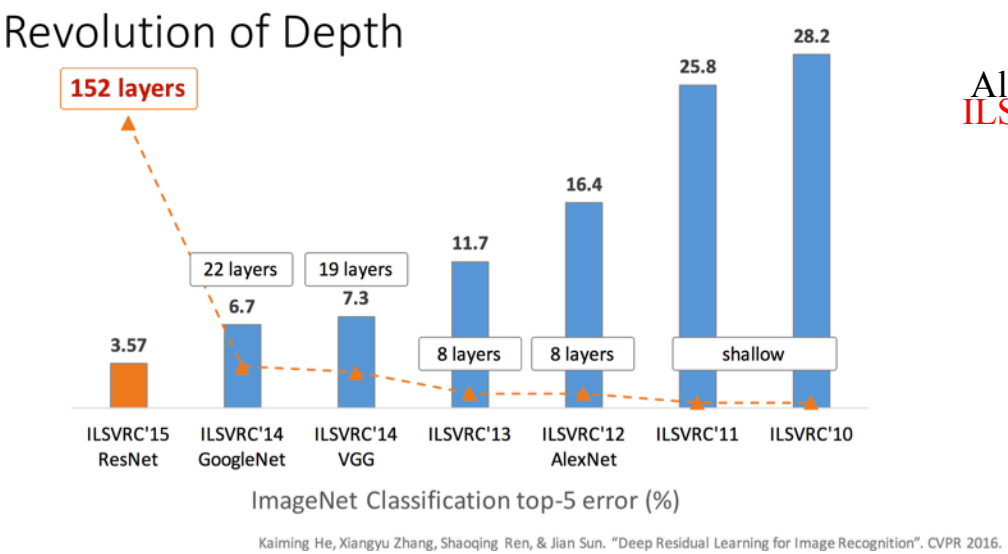
NVIDIA®

http://cvpr2017.thecvf.com/

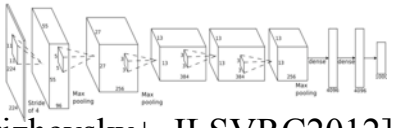
DNNの動向（4/8）

構造の深化

- 2014年頃から「構造をより深くする」ための知見が整う
- 現在（主に画像識別で）主流なのはResidual Network



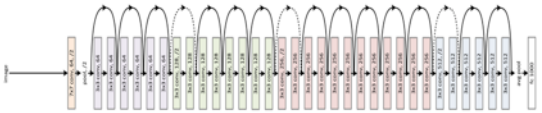
Kaiming He, Xiangyu Zhang, Shaoqing Ren, & Jian Sun. "Deep Residual Learning for Image Recognition". CVPR 2016.



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

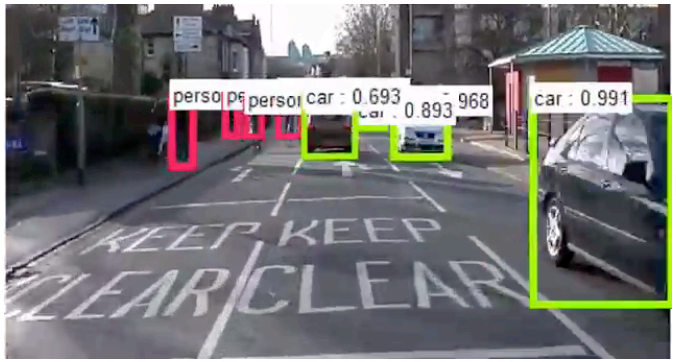


ResNet [He+, ILSVRC2015/CVPR2016]
ILSVRC2015 winner， 152層！(実験では10^3+層も)

DNNの動向（5/8）

他タスクへの応用（画像認識・動画認識）

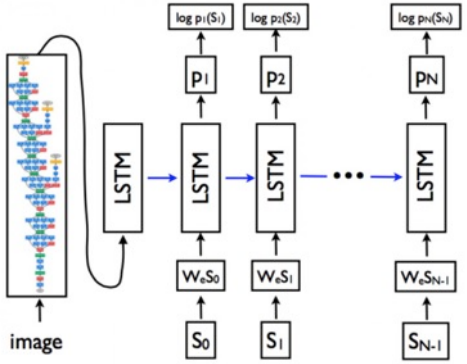
- R-CNN: 物体検出
- FCN: セマンティックセグメンテーション
- CNN+LSTM: 画像説明文
- Two-Stream CNN: 動画認識



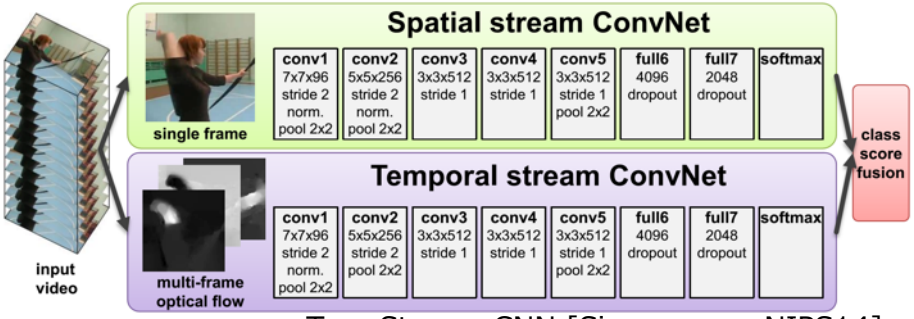
ResNet + FRCNN
[He+, CVPR16]

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

PSPNet [CVPR17]



Show and Tell [Vinyals+, CVPR15]

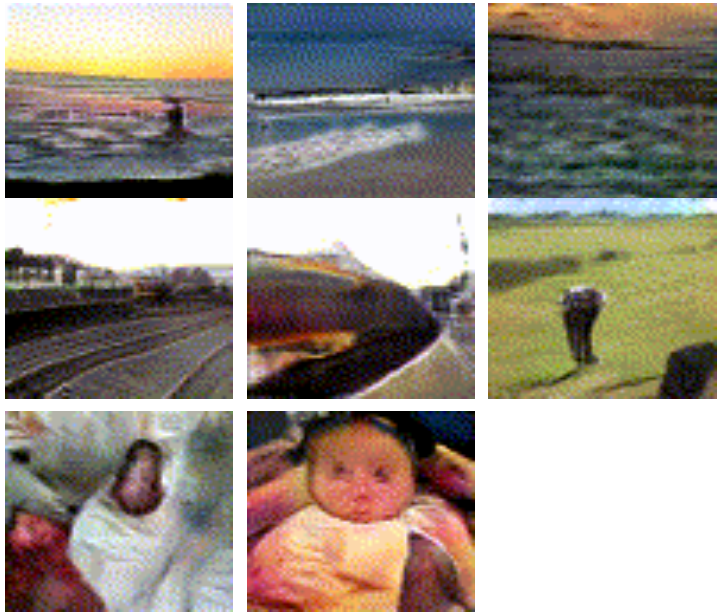


Two-Stream CNN [Simonyan+, NIPS14]

DNNの動向（6/8）

画像生成・強化学習への移行

- GAN: 敵対的学習，Generator（G）とDiscriminater（D）が競い合いリアルな画像を生成
- DQN: 強化学習モデル



Video GAN [Vondrick+, NIPS16]
http://carlvondrick.com/tinyvideo/

https://www.youtube.com/watch?v=V1eYniJ0Rnk

DNNの動向（7/8）

DNNのフレームワークが次々にリリース

– Caffe/Caffe2, Theano, Chainer, TensorFlow, Keras, Torch/PyTorch, MatConvNet, Deeplearning4j, CNTK, MxNet, Lasagne

（順不同，その他多数）

– 特に，Caffeが出てきてからCVにおけるDNNの研究は爆発的に広がった

Chainer

https://chainer.org/images/logo.png

TensorFlow

https://www.tensorflow.org/_static/images/tensorflow/logo.png

PYTORCH

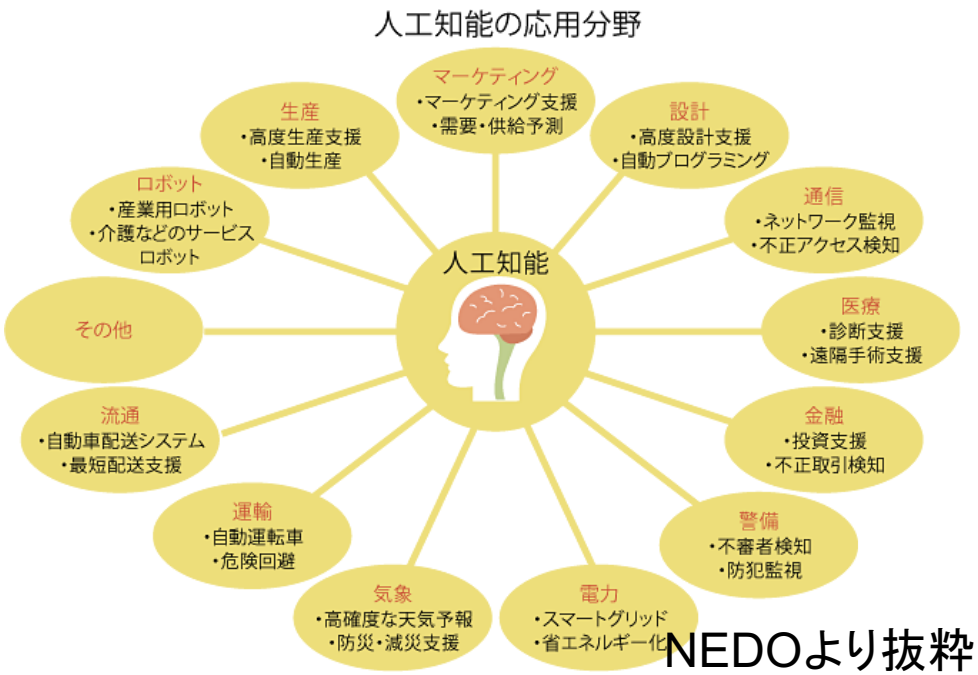
http://pytorch.org/docs/master/_static/pytorch-logo-dark.svg

日本ではChainer？ EuroではMatConvNet？ 世界的にはTensorFlow？ コード量の少なさからKerasもよく聞く
当グループではPyTorchでほぼ統一して共有（Facebookでも研究はPyTorch，プロダクトはCaffe2）

DNNの動向（8/8）

現在も進化の一途を辿り，社会実装が進む

- 自動運転/ADAS
- ロボティクス
- ファッション
- 画像/動画検索
- 物流（ピッキング等）
- 等



NEDOより抜粋

http://www.nedo.go.jp/tokushu/imgs/fn56/img_otuki01_l.png

研究者としては「こんなこともできる」を世に出したい

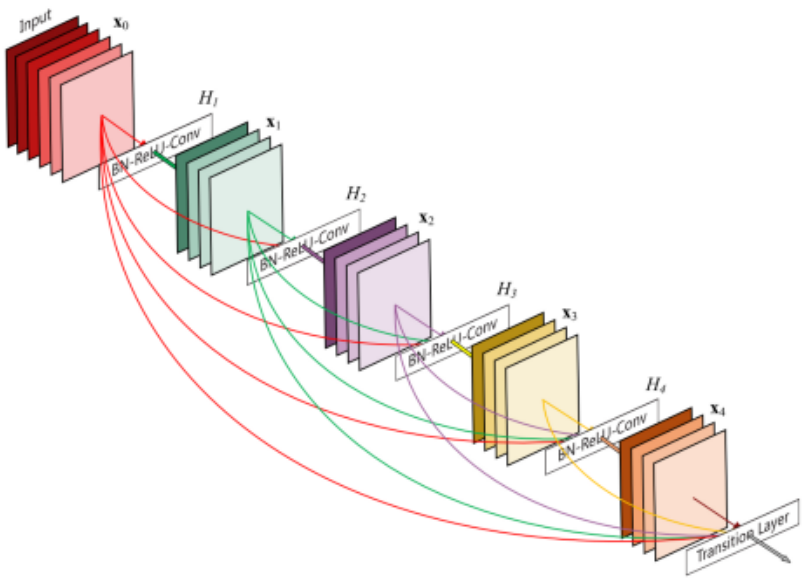
CVPR 2017の動向・気付き

今までの蓄積，今回CVPR2017論文を読んで気付いたことや海外の研究者の動向をまとめました
あまり繋がりがなく，列挙しただけですがご了承ください

CVPR2017の動向・気付き（1/19）

– （DNNは）タスク特化に加えて高精度化

- 精度を改善するための取り組み
- DNN初期からある画像識別/物体検出を例にすると…



DenseNet [Huang+, CVPR17]
CVPR Best Paper

	YOLO								YOLOv2
batch norm?		✓	✓	✓	✓	✓	✓	✓	✓
hi-res classifier?			✓	✓	✓	✓	✓	✓	✓
convolutional?				✓	✓	✓	✓	✓	✓
anchor boxes?				✓	✓				
new network?					✓	✓	✓	✓	✓
dimension priors?						✓	✓	✓	✓
location prediction?						✓	✓	✓	✓
passthrough?							✓	✓	✓
multi-scale?								✓	✓
hi-res detector?									✓
VOC2007 mAP	63.4	65.8	69.5	69.2	69.6	74.4	75.4	76.8	**78.6**

YOLO_v2 [Redmon+, CVPR17]
CVPR Honorable Mention Award

CVPR2017の動向・気付き（2/19）

– データベースの大規模化はある程度収束

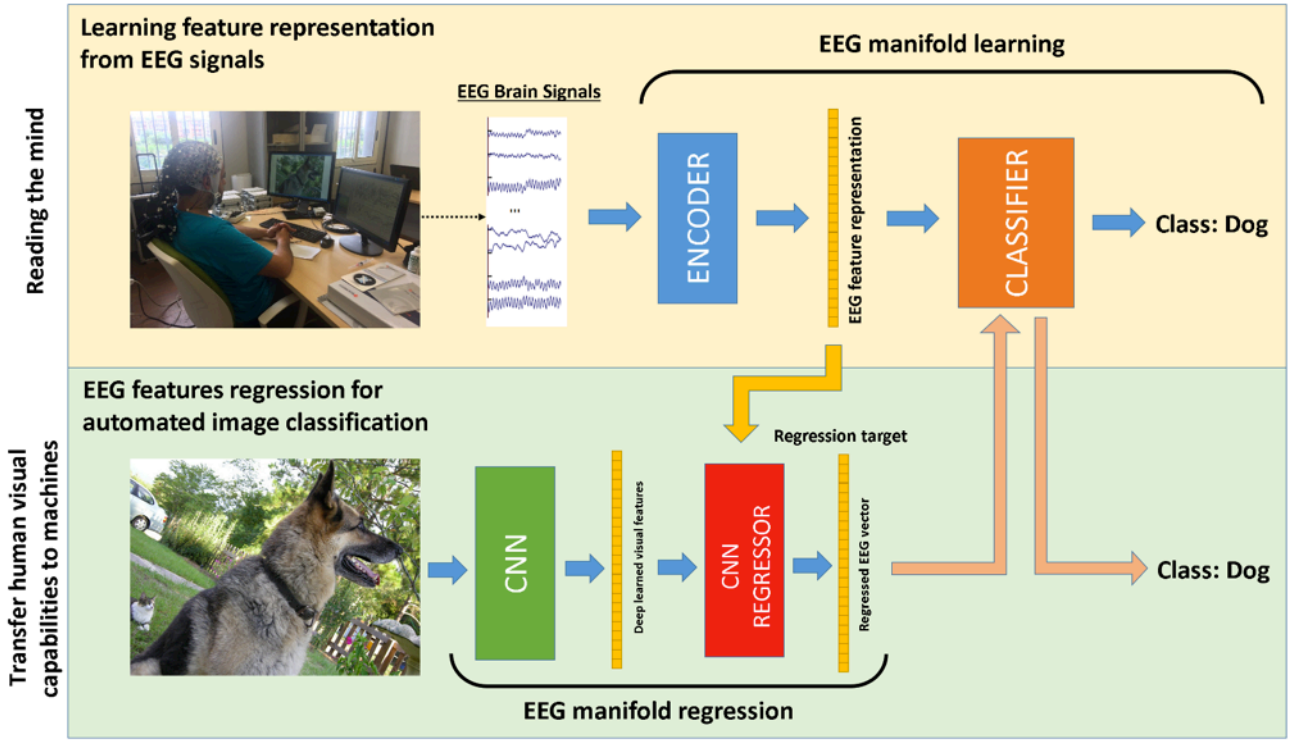
- 「ただ大きい！」だけのデータは必要ない？
- 意義のあるデータ/ クリーンなデータが重要と分かってきた
- 弱いラベル/ ラベル無し データでの手法を考案も評価高

当グループでも問題設定とともに「データを自ら作る」を徹底
画像・動画の収集，アノテーションやクロスチェックもできる限り迅速かつ大規模，網羅性を考慮

CVPR2017の動向・気付き（3/19）

– DNNの解明に向けた研究

- 中間層の挙動解明（e.g. Network Dissection）
- クラス間の境界を探る（e.g. Perturbation）
- 脳の信号から画像識別する研究も見られた（Deep Learning Human Mind; 下図）



CVPR2017の動向・気付き（4/19）

– 大規模な学習を実現するための研究
  - Weakly Supervised Learningによる比較的容易に取得可能な教師データ（画像単位，クリックなど）からの学習（21件）
  - CGなどから人工的な学習データを生成 (静止画，動画問わず)
  - より実データに近い人工データをGANなどにより生成
  - ラベル無しデータにラベルを付与しながら学習 (Self-supervised)

CVPR2017の動向・気付き（5/19）

– 従来のタスクをベースに複雑化

- 多人数を対象にリアルタイム処理
- より多様（姿勢，環境…）で実世界に近いデータを対象
- 複数のタスクを単一のフレームワークで処理

新しいデータセットやアノテーションを用意して
少しでも独自の設定にしようという工夫

CVPR2017の動向・気付き（6/19）

– ボリュームデータへの移行

- ボリュームデータとは，画像xy空間から新しい次元を追加
- 奥行Zを含めた実空間（XYZ），時間軸Tを含めた動画像（XYT）等
- 基礎的なデータや手法の整備をまずは進めている

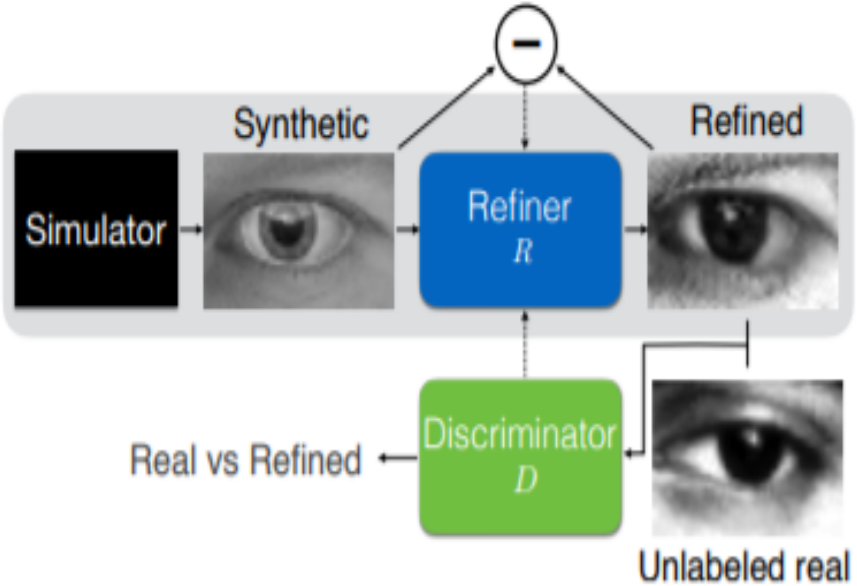
Figure 1. Example reconstructed spaces in ScanNet annotated with instance-level object category labels through our crowdsourced annotation framework.

ACTIVITYNET	Kinetics	SPORTS-1M	YouTube 8M
6766 / 13220	28108	306245	1M / 8M
51 / 101	200	400	487 / 4716

CVPR2017の動向・気付き（7/19）

– GAN/DQNの効果的な利用

- Apple: 合成データ（Synthetic Data）をリアルにする仕組みを導入してリアル画像と見分けがつかない画像へと変換する（Best Paper）
- Google: ドメイン変換をGANで実現
- Snap&Google: 画像説明文のパラメータを強化学習で強くする
- アイディア次第



CVPR2017の動向・気付き（8/19）

– 根強く残る分野へ攻める

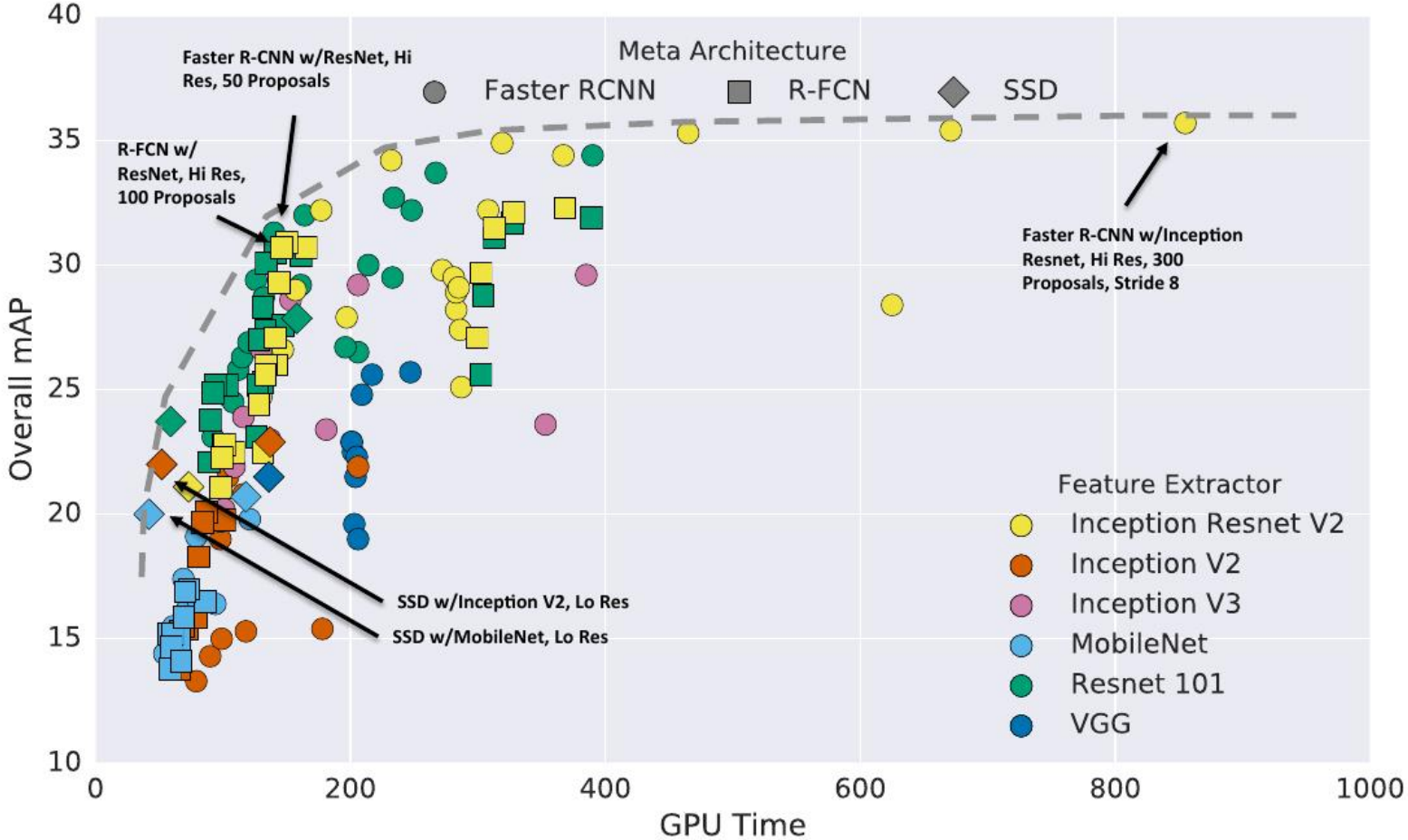
- Computational Photography，カメラ幾何等の純CV
- まだDeep Learningが入り込んでいない（というか入らなくても良い分野ももちろんある！）
- Machine Learningのセッションと比較して人数が少ない（= 査読を突破する可能性も上がる？）

今の時代，（逆に）DNNをやらない研究の方が通しやすい？

CVPR2017の動向・気付き（9/19）

– 徹底的な比較論文は通る

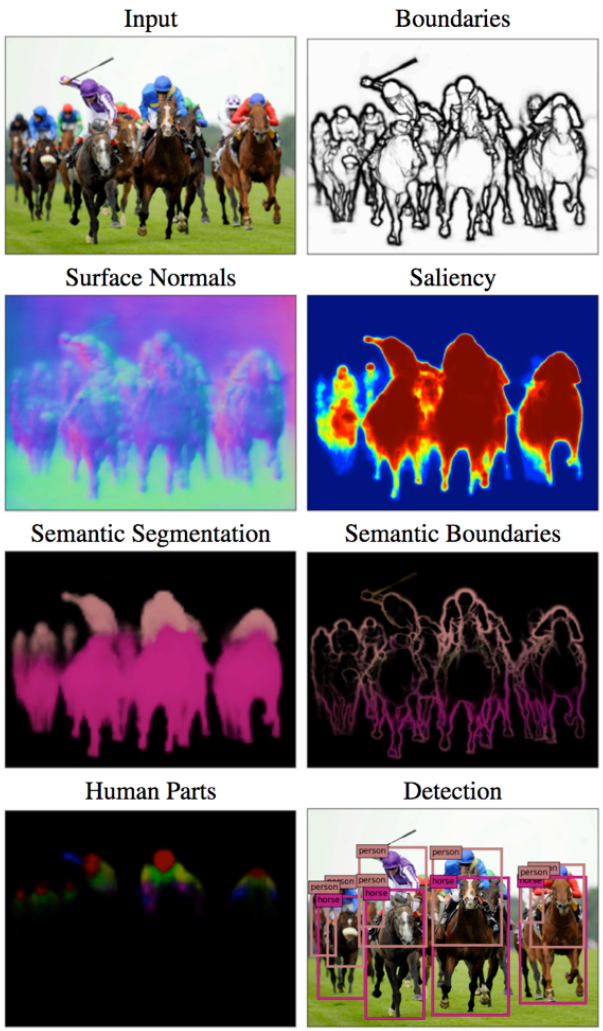
- Speed/Accuracy Trade-off（Google）
- HPPatch: 局所特徴量のベンチマークと徹底比較

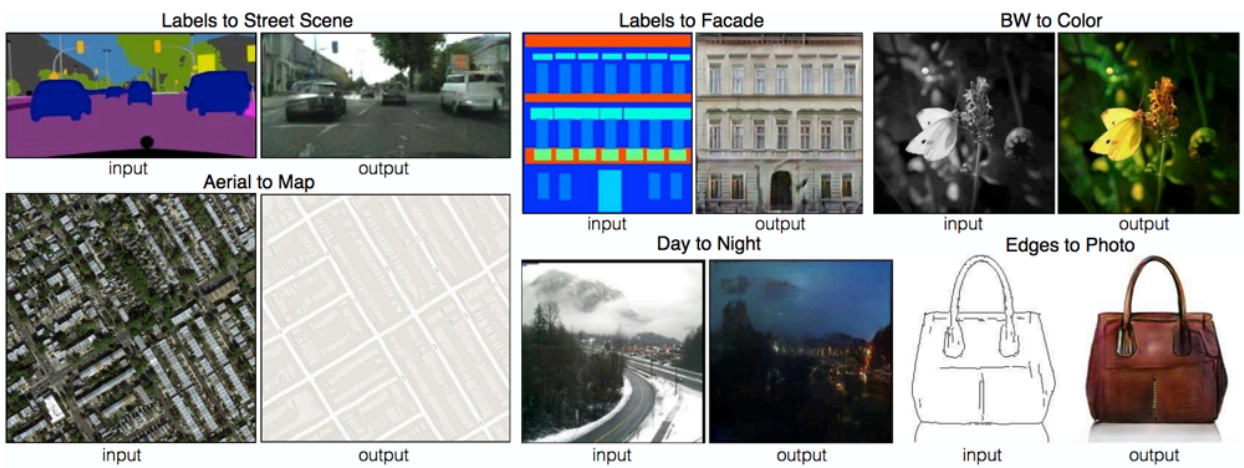


CVPR2017の動向・気付き（10/19）

– ある意味万能・多機能なネットワークを構成

- UNet, pix2pix





CVPR2017の動向・気付き（11/19）

– マルチモーダル・マルチタスクを扱うDNNが増加

- DNNが縦・横に伸びていく
  - ResNetのSkippingConnectionの成功を受けて？
  - GPUの高性能化（特にメモリ増大）の影響大？
- データセットの多様化
  - 応用的課題に使えるデータセットの出現を受けた研究が出現
- 異なる機能のモジュールを結合しても十分使える
  - タイトルを「Joint」で検索すると…
    - » <u>CVPR2017 21件</u>
    - » CVPR2016 16件
    - » CVPR2015 17件
    - » CVPR2014 15件
- DNNの多分野利用の加速

CVPR2017の動向・気付き（12/19）

- Oral, Spotlight, Posterのどれでも，十中八九，DNNの構造をいじってstate-of-the-artという新規性の主張
  - 手法的新規性の重視
  - 何やってもうまくいくので手法的新規性が出しやすいせい
  - DNNの構造が最大のハイパーパラメータ
- 既存技術の性質の調査もマイナーながらある
  - 問題設定，論文の書き方がうまい？ =ハードルが高い？

CVPR2017の動向・気付き（13/19）

– 事前に論文の宣伝を行うことも重要？
  - 毎回通している人に聞いてもarXivに載せることが重要と話している（うまく活用している）
  - 今回，YOLO9000やpix2pixの戦略が参考になった（真似できるかどうかは置いといて。。。）
    – 何れもコード公開，arXivへの論文投稿はもちろん，Twitterでも効果的に拡散させていた
    – 特に，YOLO9000の圧倒的な開発力は評価に価するし，ポスター（下図）はズルい

https://pbs.twimg.com/media/DFnnqn0VwAExuiU.jpg

CVPR2017の動向・気付き（14/19）

– アイディアや研究テーマが似通ってきた？
  - 同データ（ImageNet/PascalVOC等）で同手法（ResNets/R-CNN/FCN等）をベースにすると必然的にアイディアは重なる
  - （重要１）完成度を高めないと取り組み自体消される
  - （重要２）問題設定・手法ともにBrave Newなアイディアが重宝

CVPR2017の動向・気付き（15/19）

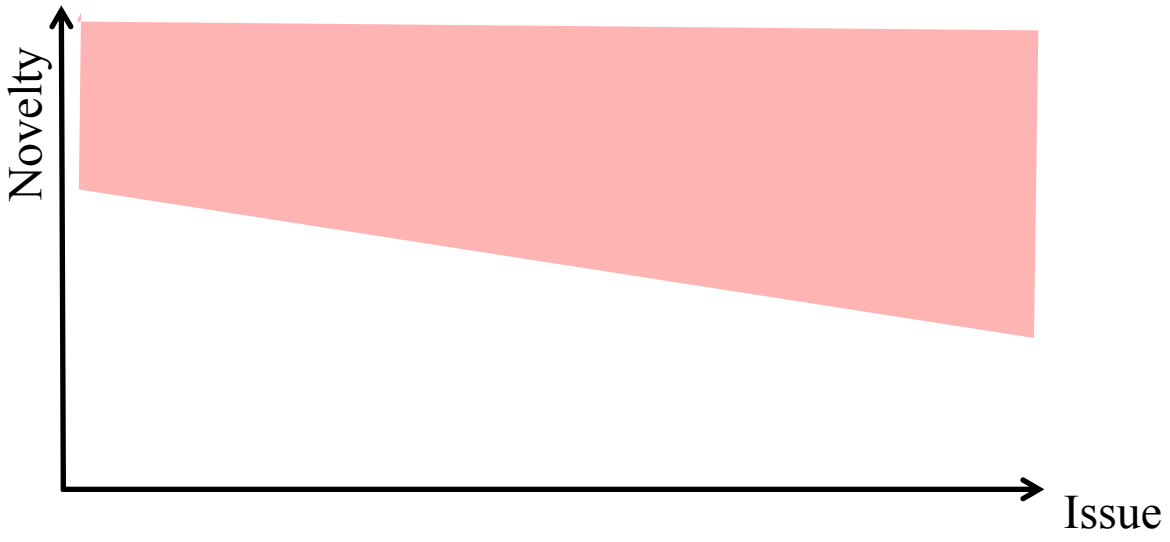
– 思いついたアイディアはその年のうちにやる
  - 次の年に同じような研究が出てくる（可能性が高い）
  - 我々のグループ内の（ブレストにより出てきた）アイディアが今年のCVPRでも散見された
  - 誰でも思いつきそうはことは世界中で同時多発的に研究される

＃「せっかく思いついていたのに」は意味なくて，実際にやって見せないと意味がない

CVPR2017の動向・気付き（16/19）

– どんな論文が通っている？

- （当たり前だが）手法として新規性のある手法が通る
- データセットだけを提案する論文はよほど面白みがないと通らない（らしい）
- 問題設定をIssue，手法をNoveltyとすると（直感的に）下記エリア
  - Novelty重視



CVPR2017の動向・気付き（17/19）

– 全部読むより現場に行こう！

- 当グループは2015年，CVPR論文を全て読んだが，

「現場に来て情報収集」する方が良い！

- さらに言うと複数人で来て議論を行う
  - 情報収集を共有（複数人で来て自然に議論している研究グループは強い）
  - 論文の著者と直接話す
  - 欲を言えばその場で研究戦略やテーマを考案

CVPR2017の動向・気付き（18/19）

– 分野間がシームレスになる傾向？
  - これまでよりもアイディアの重なりが大きい（CVで戦うツラさ）
  - DNNがツールとして整い他分野の研究者がCVPRに投稿

  => 逆を言うとDNNを武器に他分野で発表するチャンス？

CVPR2017の動向・気付き（19/19）

– そろそろ次の時代へ？

- 生の声を聞いているとDNNそろそろ飽きたという声もちらほら
  - パラメータ調整問題
  - 複雑ネットワーク構築問題
  - 大規模データセットアノテーション問題　等
- （去年くらいから）Post-DNNを探している研究グループもあるのでは？

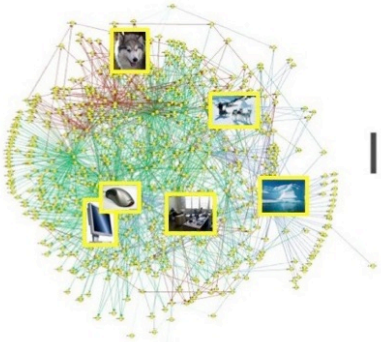
CVPR2017の動向・気付き（Bonus slide）

– 世界トップ研究室の戦略

- 問題意識を先に論文にする
- DBを提案したラボを例にとると。。。

Fei-Fei Li
(Stanford Univ.)
http://vision.stanford.edu/img/feifei2.jpeg

Antonio Torralba
(MIT)
http://web.mit.edu/torralba/www/assets/images/antonioTorralbaS.jpg



IMAGENET

CVPR2017の動向・気付き（Bonus slide）

– データセットを提供した後に何をするか？

- コンペティションを開催する（問題を解く方ではなく, 出す方に回る）
  - 与えられた問題の精度を高くできる人はいくらでもいる
  - 出てきた手法はツールとして使っている
- 自分たちはその間データを拡張および次の問題を創る研究
  - イケメン研究者になろう！

問題を創るために（徹底して）論文を読む，アイディアを考える，手を動かす

cvpaper.challenge は何を目指しているか？

cvpaper.challengeの目指すもの

次世代のCV研究を創ってみせる

- 問題がさき，解法はあとの{Concept-, Vision-}Driven
- どうやってモデルを改善するか？ではなく今の技術+αで何ができるか？を考える

cvpaper.challengeの目指すもの

技術で面白いことをやる・見せる

- こんな面白いことができる！と未来を見せる論文
- 今の技術を「使い，組み合わせ，洗練させ」ギリギリ実現可能な将来ビジョンを創り出す
- 新しい技術はどんどん使うが，その代わりそれでどんな未来ができるのかを想像して研究する

Transitional Action Recognition

BMVC2016

遷移状態を把握して行動予測

H. Kataoka et al. “Recognition of Transitional Action for Short-Term Action Prediction using Discriminative Temporal CNN Feature” in BMVC, 2016.
(Acceprance rate: 39%)

Human Action Recognition without Human

ECCVW2016

背景領域が人物行動認識に与える影響を評価

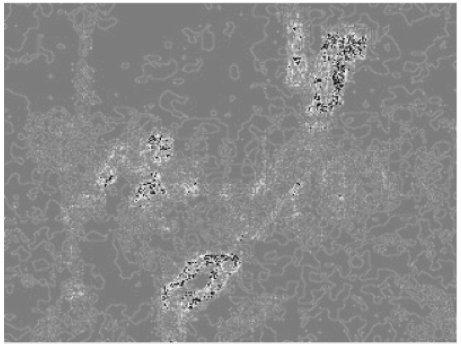
H. Kataoka et al. “Human Action Recognition without Human” ECCV 2016 Workshop Oral (Best Paper) & MIRU 2017 Oral.

Accelerated Images

ECCVW2016

加速度を2D画像に投影して特徴抽出

- 位置 (RGB), 速度 (Flow)に対して加速度を追加
- 特徴抽出はDNNに任せる



Physics quantity	I	I' (1st-order diff.)	I'' (2nd-order diff.)
Input	RGB	Flow image	Acceleration image
Stream	Spatial-stream	Temporal-stream	Acceleration-stream

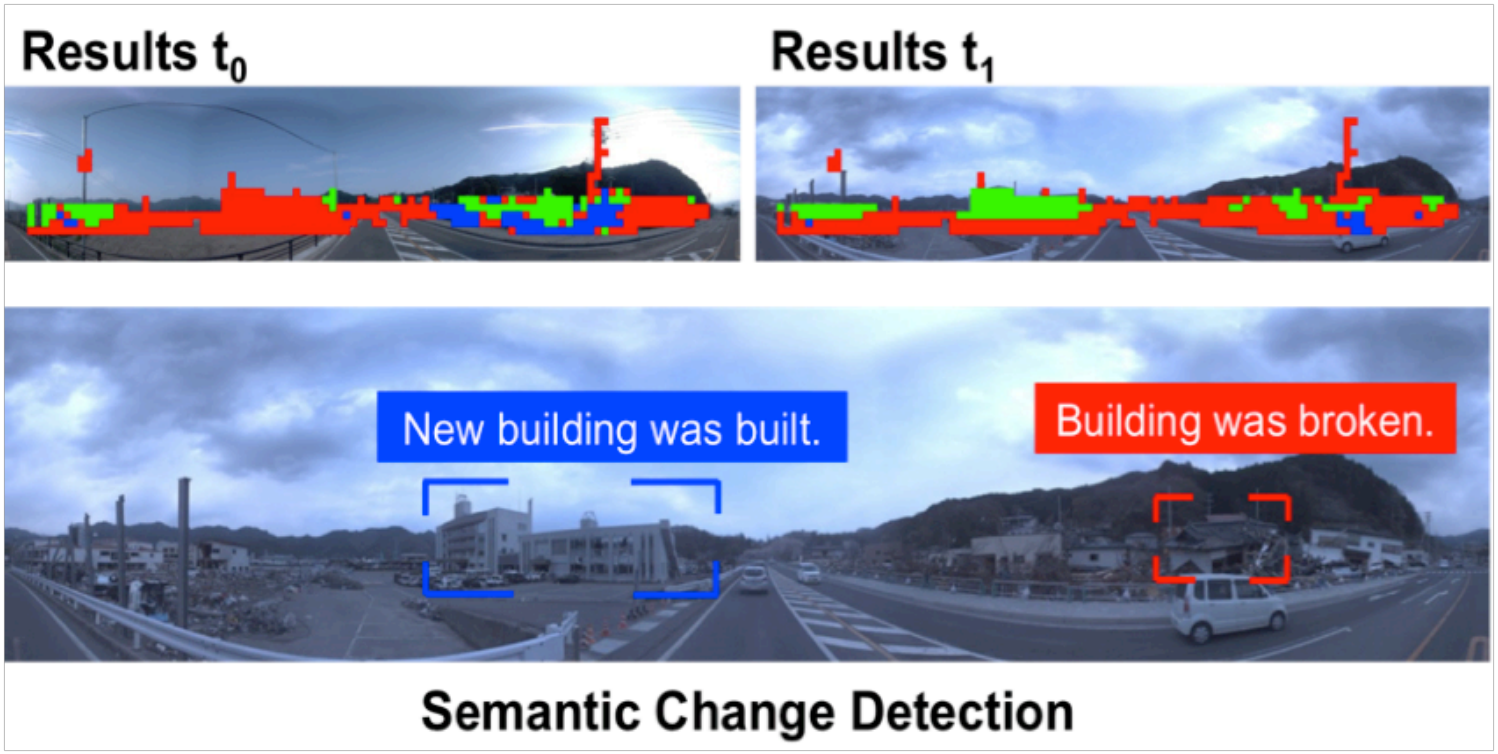
H. Kataoka et al. "Motion Representation with Acceleration Images" in ECCV 2016 Workshop, 2016.

Semantic Change Detection

arXiv2016

意味的な変化を捉える

- 直感的には「変化検出」+「セマンティックセグメンテーション」
- 「何がどのように」変化したかを理解



H. Kataoka et al. “Semantic Change Detection with Hypermaps” in arXiv 1604:7513, 2016.

Movie Poster Database (MPDB)

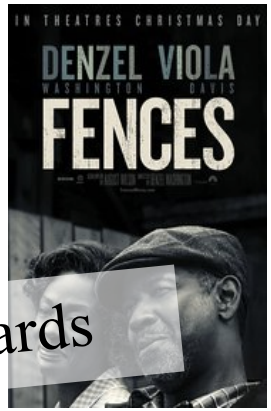
MVA2017

映画ポスターのみからアカデミー賞を予測

Let's try!

- Can you correct the Academy Award 2017?
- Which movie poster do you like, and why?

Nominated movies in 89th Academy Awards



Y. Matsuzaki et al. "Could you guess an interesting movie from the posters?" in MVA, 2017.

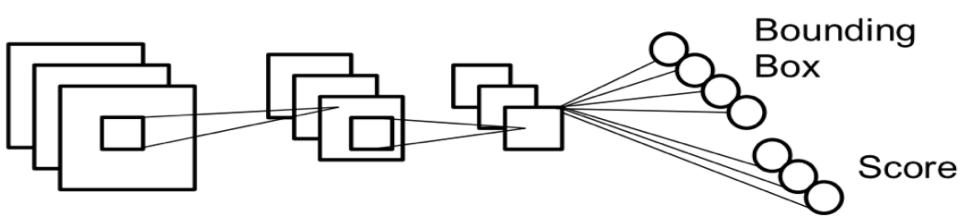
Pedestrian Near-Miss Detection

MVA2017

歩行者のニアミス状態を検出

- 通常状態とニアミス状態の歩行者の動作や姿勢は異なる
- ニアミス位置やその危険度も含めて検出可能

Input (224 × 224)



CNNs

Fully Connected

Output Example

Risk high

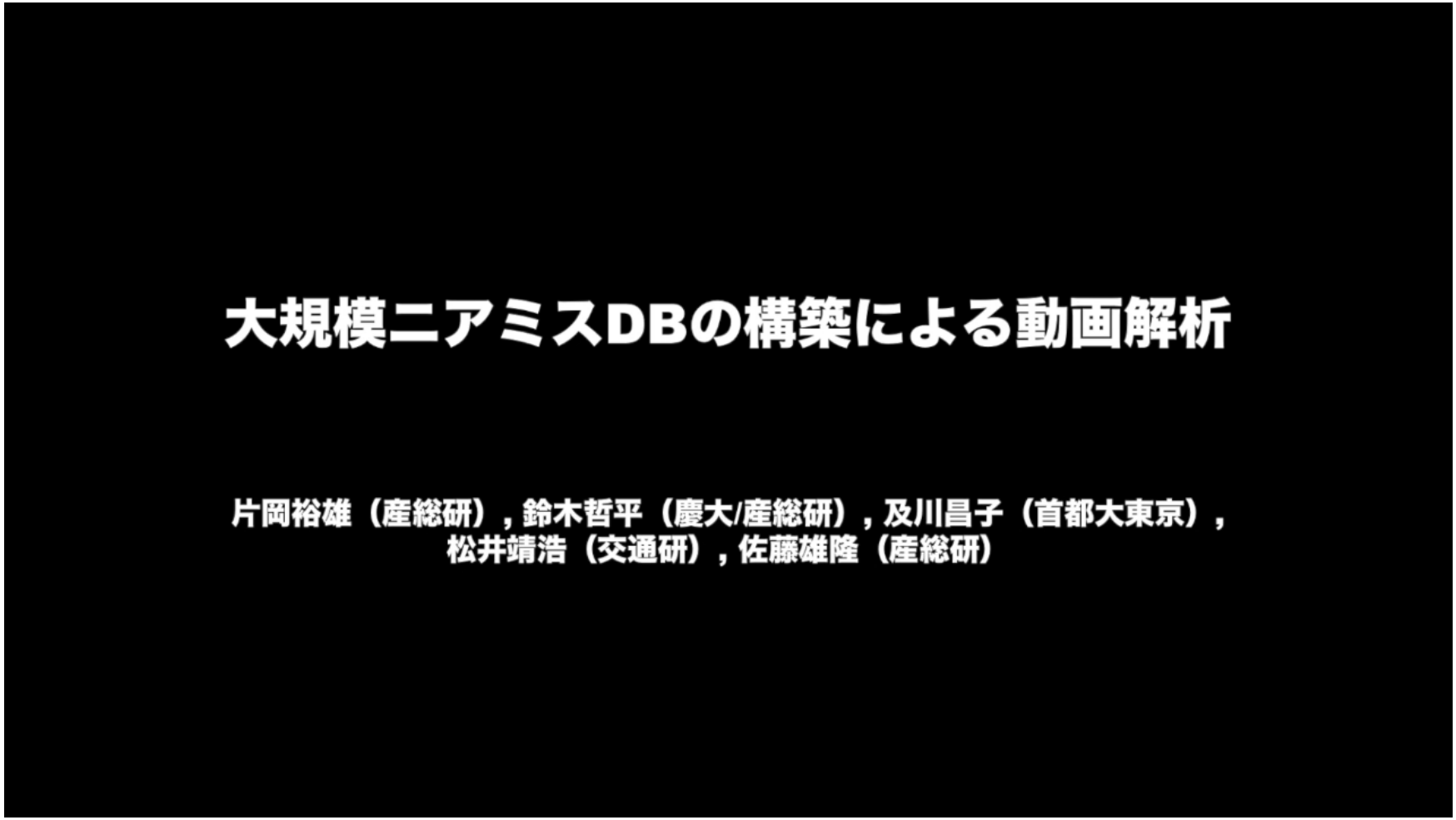
Risk low

T. Suzuki et al. “Pedestrian Near-Miss Analysis on Vehicle-Mounted Driving Recorders” in MVA, 2017.

Near-miss Incident DB (NIDB)

SSII2017

世界初の動画認識に向けた大規模ニアミスDB



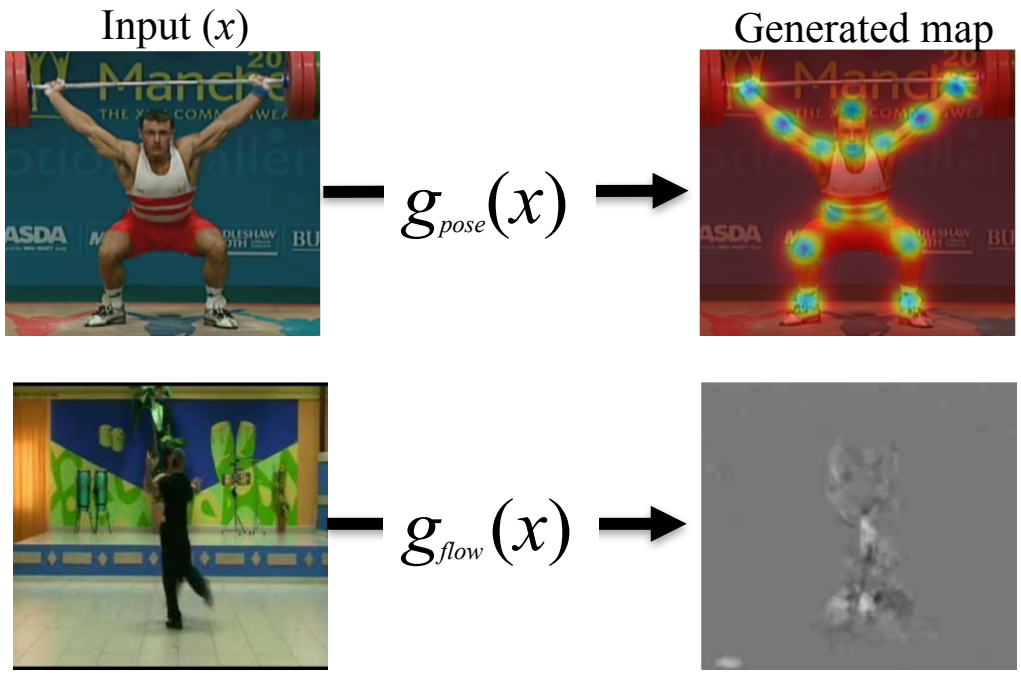
H. Kataoka et al. “Drive Video Analysis for the Detection of Traffic Near-miss Incidents” in SSII, 2017.

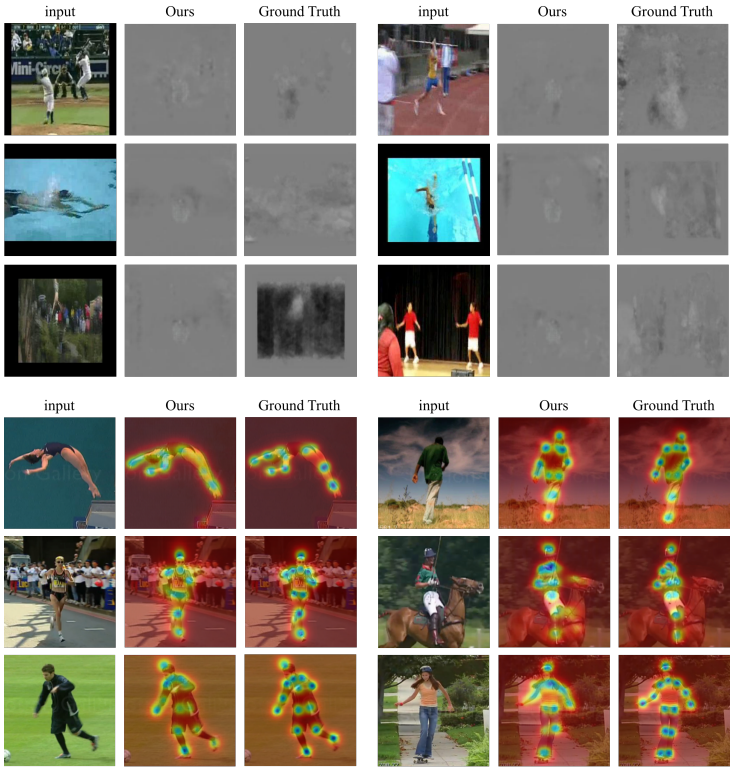
Generated Motion Maps

CVPRW2017

入力から直接 動作解析に必要なマップに変換

- フレームワークにはpix2pixを使用
- Poseのヒートマップ，Stacked Flow画像を生成





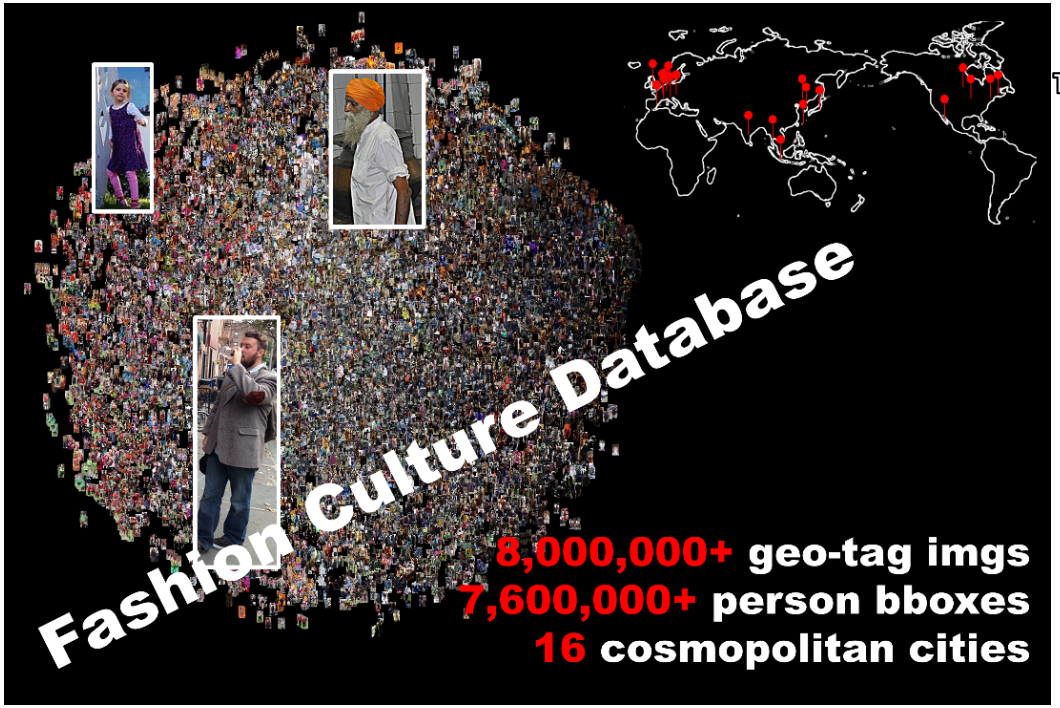
Y. Matsuzaki et al. “Generated Motion Maps” in CVPRW, 2017.

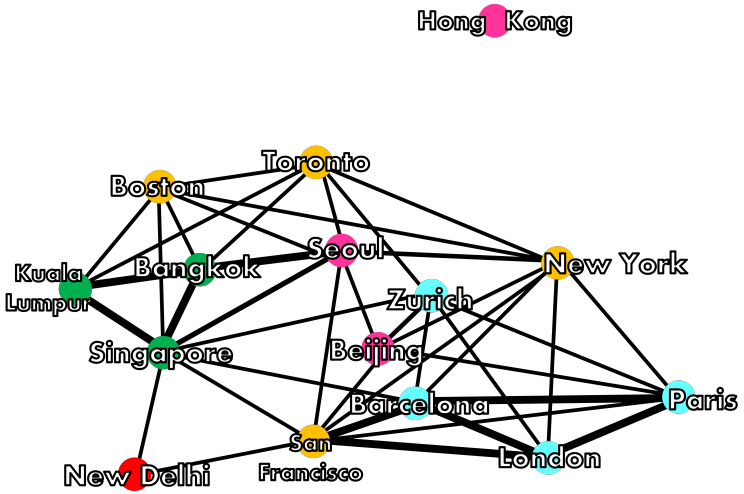
Dynamic Fashion Cultures

MIRU2017

大規模DB（76M）からファッショントレンドを発見

- 動的に変化するファッション（服装）を捉える
- 世界主要都市の動向をSNS画像から推定





2014 **2015** **2016**

K. Abe et al. “Dynamic Fashion Cultures”, arXiv, 2017 & MIRU 2017 Oral. (Best Student Paper)

3D ResNets

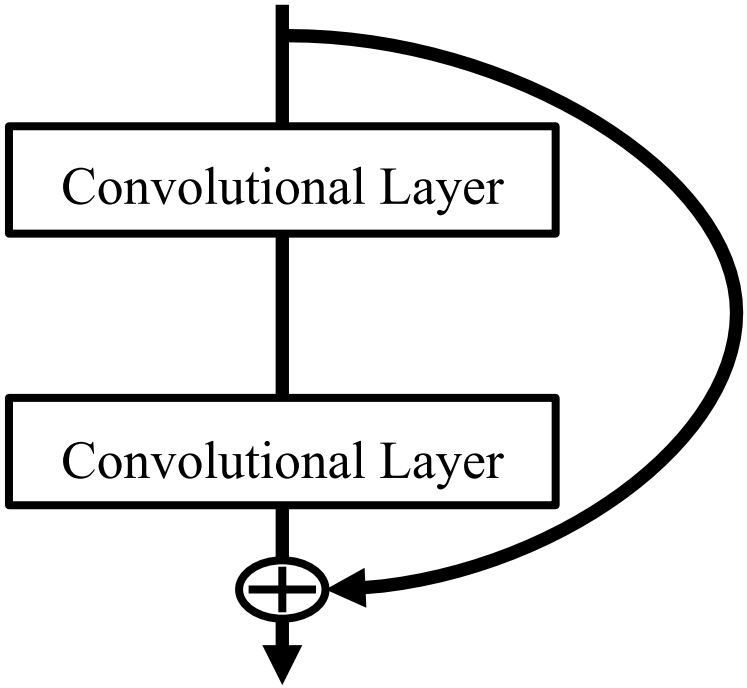
ICCVW2017

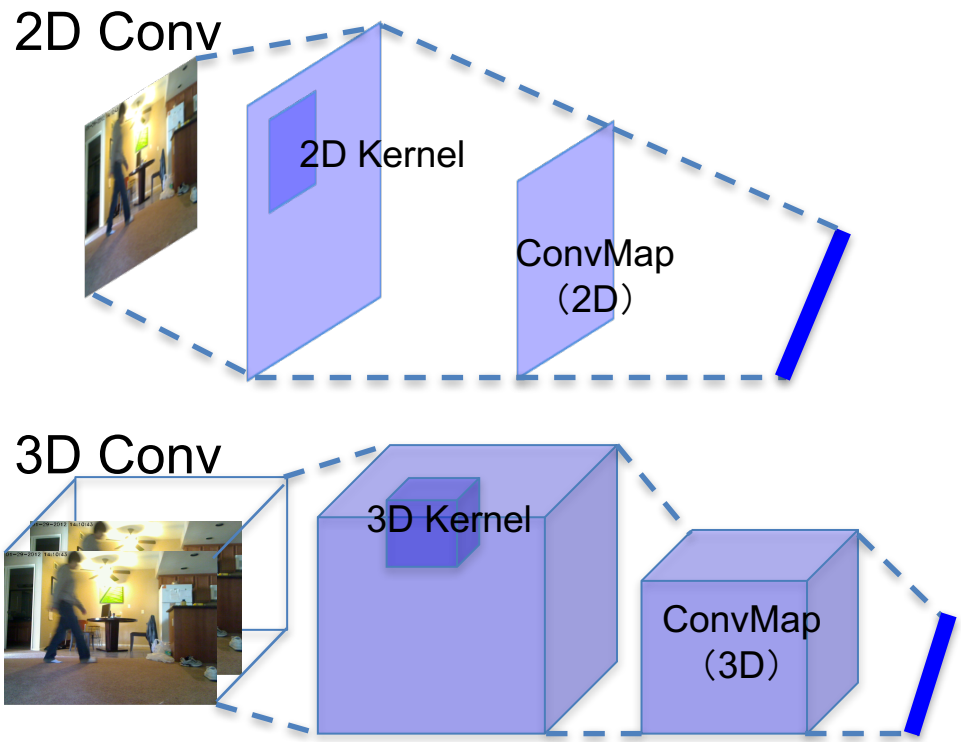
ResNetsのカーネルに3D Convを使用

– 動画認識用（xyt）のResNetsを構築

– [公開済み] Kinetics Dataset Pre-trained Model

https://github.com/kenshohara/3D-ResNets





K. Hara et al. “Learning Spatio-Temporal Features with 3D Residual Networks for Action Recognition”, ICCVW, 2017.

（コード&モデル公開済）

Fractal DB

2017

完全自動で学習済みモデルを生成

– 自然界の特徴を参考に強い表現学習ができないか？

Y. Matsuzaki et al. “FractalDB” in preparation.

コンピュータビジョンの現状

独創的な問題 and/or 強い手法がないと記録・記憶に残らない

- トップ会議の論文とて例外でない（CVPRは今年783本）
- 中途半端に分割した複数論文よりもパーフェクトな１本
- 動画やスライド公開・コード共有・DBリリースなども（できる限り）徹底して揃える

他の追随を許さぬ強い手法を作る！

独創的な問題！強い手法！

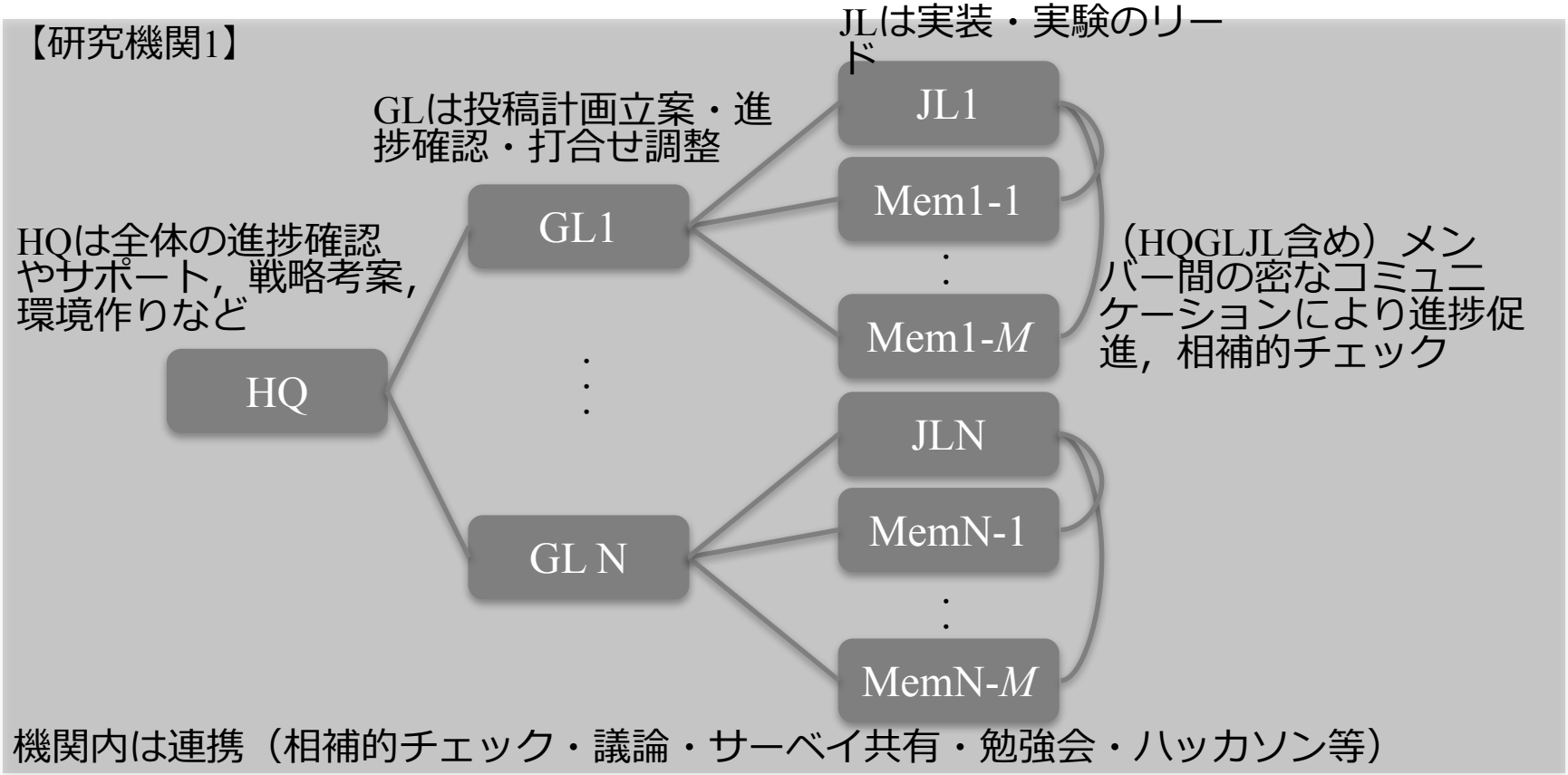
- 受賞論文や注目論文にあるように「極めて高速かつ高精度」を実現，コードをリリースして分野に貢献が理想
- 問題設定に対しての強い手法でも構わない
- 作れたらいいですね！=> e.g. DenseNets, YOLO9000, PAF, PSPNet…

研究体制への考え方

今まで以上にチームの力が重要

- 高品質論文（前スライドより）には１人のパワーでは不十分？
- 当グループでは仕組みを再考
  - 通常の学生：１人１テーマ３年間継続（学部~修士を想定）
  - cvpaper.challenge：２～４人１テーマ１年でテーマ拡張/変更

どうやって研究しているか？



- 知識：論文・Web・講演資料・サーベイテンプレート等から把握
- テーマ：メンバー間の議論により洗練
- 進捗：打ち合わせにより把握
- コミュニケーション：ITツールで補助 (e.g. Slack, Twitter, FB, LINE)
- 論文：多重チェックが重要，即ち早め（テーマ立案・実験と同時）に書く

研究方針

CV系に限らず他分野へ投稿する

- （当然だが）CVに拘らない
- 分野の成熟は他の分野の手助けになるということ

外の世界に出て行こう！

研究方針

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を復活
- 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
- もちろん，今の技術を効果的にアレンジ

OUR GOAL

- コンピュータビジョンや周辺分野にて国際的に競争力のあるグループを作る
  - トップ会議に年間１０本通るようなチーム

- 社会実装を促進するための枠組み・仕組みを持つ
  - 研究した技術を社会に還元する「強い連携」

- 知識を共有・浸透させるためのプラットフォームを拡大
  - サーベイを始めとした最新の知識を素早く提供

Read 1,000+, Write 10+ papers.

JOIN US!

- １年で３年分の業績を出したい人！研究者として劇的に成長したい人！

- 産総研（AIST）のメンバーとして参加したい！
  ー 学生はRA/インターン，企業の人も技術研修員として可

- コラボレータとして参加する！
  ー 個人でもグループでも可です，クレイジーなアイディアを実現しましょう

- 論文を読んで知識を共有する！
  ー 読むだけでも歓迎！（企業の人はコレが多い）

- ただ友達になりたい！
  ー いつでも歓迎！

コンピュータビジョンの今を映し未来を創るチャレンジ,
一緒に挑戦しませんか？