

人工知能技術戦略会議等主催  
第1回AIチャレンジコンテスト  
表彰式

## プレゼンター（敬称略、五十音順）

---

安西 祐一郎

人工知能技術戦略会議 議長、独立行政法人日本学術振興会 理事長

大崎 真孝

エヌビディア合同会社 日本代表 兼 米国本社副社長

大淵 栄作

株式会社デジタルメディアプロフェッショナル 常務取締役開発統括部長

齊藤 秀

株式会社オプト 最高解析責任者CAOデータサイエンスラボ代表

成田 一生

クックパッド株式会社 最高技術責任者

宮内 淑子

株式会社ワイ・ネット 代表取締役社長

# AIチャレンジコンテスト 概要

## ■趣旨

国内の学生・社会人を幅広く対象として、先端的な人工知能技術の開発とビッグデータ活用の能力を競う場として、「AIチャレンジコンテスト」を開催。

人工知能技術を開発し活用できる人材の発掘や、実際の課題・データを対象とした研究開発や優れた参加者の技術・アイデア等から波及する人材育成効果を期待。

## ■テーマ

料理画像jpgデータからの、①料理の領域の検出 ②料理の種類のカテゴリ

## ■データ

10万枚の料理画像データ（クックパッド社御提供）

（開催主体）

・主催 人工知能技術戦略会議、内閣府、文部科学省

・協賛     

・後援 総務省、経済産業省、情報通信研究機構、理化学研究所、産業技術総合研究所、科学技術振興機構、新エネルギー・産業技術総合開発機構、情報処理学会、電子情報通信学会(パターン認識・メディア理解研究会)、電子情報通信学会(食メディア研究会)

# コンテスト概要

部門	①料理領域検出部門	②料理分類部門			
課題	各画像における料理領域の特定	各画像の料理の分類			
データ	3万枚の料理画像と料理領域の情報	7万枚の料理画像と25種の料理カテゴリの情報			
賞	<b>人工知能技術戦略会議 議長賞（入賞者中の特に優れた人1名を審査）</b>				
	<b>最高精度賞</b> 1位：賞金20万円（DMP社提供） 2,3位：nintendo New3DS-LL（DMP社提供） 4,5位：nintendo 2DS（DMP社提供） 6-10位：ノベルティ（各社提供）	<b>アイデア賞</b> 1位：賞金10万円（クックパッド社提供）+ノベルティ（各社提供） 2,3位：GeForce GTX 1060 6GB（NVIDIA社提供） 4,5位：GeForce GTX 1060 3GB（NVIDIA社提供）	<b>最高精度賞</b> 1位：NVIDIA TITAN X（NVIDIA社提供） 2,3位：nintendo New3DS-LL（DMP社提供） 4,5位：nintendo 2DS（DMP社提供） 6-10位：ノベルティ（各社提供）	<b>アイデア賞</b> 1位：賞金10万円（クックパッド社提供）+ノベルティ（各社提供） 2,3位：GeForce GTX 1060 6GB（NVIDIA社提供） 4,5位：GeForce GTX 1060 3GB（NVIDIA社提供）	
	提出物	予測結果	予測結果 +レポート	予測結果	予測結果 +レポート
	評価	精度上位10名	基準精度以上の優れた5名を審査	精度上位10名	基準精度以上の優れた5名を審査

# 審査員9名（敬称略、五十音順）

## ○ 安西 祐一郎 氏

人工知能技術戦略会議 議長、独立行政法人日本学術振興会 理事長

## 井崎 武士 氏

エヌビディア合同会社 ディープラーニング部 部長

## 大淵 栄作 氏

株式会社デジタルメディアプロフェッショナル 常務取締役開発統括部長

## 齊藤 秀 氏

株式会社オプト 最高解析責任者CAOデータサイエンスラボ代表

## 成田 一生 氏

クックパッド株式会社 最高技術責任者

## 服部 幸應 氏

学校法人服部学園 服部栄養専門学校 理事長・校長

## 宮内 淑子 氏

株式会社ワイ・ネット 代表取締役社長

## 山田 誠二 氏

国立情報学研究所 教授、人工知能学会 会長

## 大和 喜一 氏

株式会社IDCフロンティア 技術開発本部 R&D室 主任研究員

# 課題内容 (①料理領域検出部門)

## ■データ

学習画像 : 2万枚

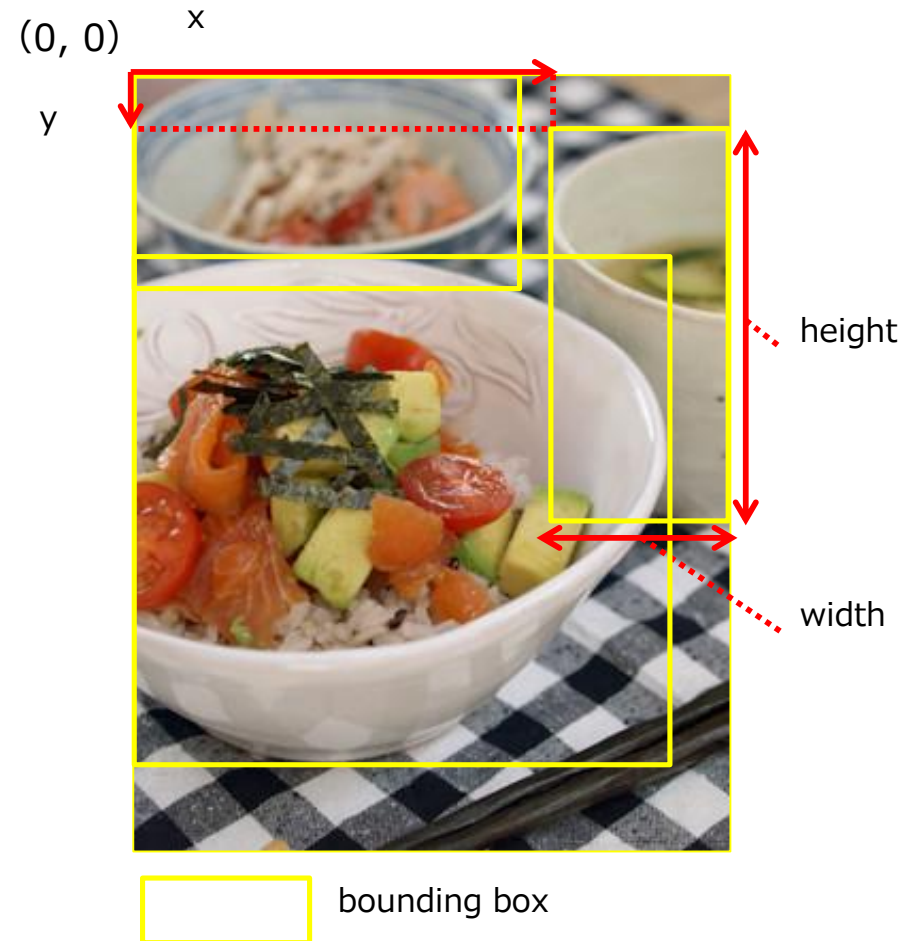
評価画像 : 1万枚

## ■課題

画像に対して、料理が写っている領域 (1つ以上) を、bounding box = (x, y, width, height) として割り当てる。

## ■評価関数

mean average precision (MAP) を使用。



# 課題内容 (②料理分類部門)

## ■データ

学習画像 (ラベル付き) : 1万枚

学習画像 (ラベル無し) : 5万枚

評価画像 : 1万枚

## ■課題

画像に対して、25種類のカテゴリ (サンドイッチ、スライSPAN、素麺、うどん、クリームパスタ、トマトパスタ、カレー、リゾット、寿司、味噌汁、ポタージュスープ等) の中から、1つカテゴリを付与する。



コンテストの総参加者数：

**631** 名

予測結果の総提出件数：

**2732** 件

(開催期間：平成29年1月10日～3月31日の  
12週間、81日間)



# アンケート結果

## ■ 作業時間

平均：62.75時間

最大：230時間

## ■ 感想

- とてもおもしろいコンテストで、**大量な実データが入手できることやほかの方と公平にスコア比較できることが、勉強と研究のモチベーション**につながります。今後もぜひ開催してください。計算資源で差がつくコンテストだと思いました。楽しかったです。このような機会を提供していただきありがとうございます。
- Cookpadの画像データという**現実世界のデータを解析でき、とてもいい経験になりました！**
- 面白いテーマがあれば**今後もぜひ参加させていただきたい**と思います。こういった機械学習系のコンテストが日本でももっと盛んになると良いなと思います。
- もう少しコンテスト期間が長いとよかったかも。画像認識は既に手法がかなり確立されてきているタスクだと思いますが、**ラベルなしの訓練データが明示的に与えられているところが工夫のしどころとして面白い**と思いました。色々大変でしたが、とても勉強になったと思います。
- 事前にCookpadのMeetUpにて、コンテストルールの疑問点を聞いたのが、今回のアイデアに繋がりました。**Deep Learningでの画像認識は初めての経験だったのですが、画像認識部門は典型的かつ良質なデータだったこともあり初心者にも取り組みやすい課題**でした。
- **問題設定が非常に実践的**で面白かったです。また勉強してチャレンジしたいです。最終日に過去の高精度のやつを出してスコアがあがったり、頑張りで精度が上がったりなど、スコアはあくまで参考なのだと思いました。コンテスト形式で同じ課題を他の方と競争する形でやるのは、ここまではいけるはずという目標がはっきりするので、チャレンジしがいがある。**最初はpythonのコードも殆ど分からない素人でしたが、コンテストを通して多くを学習できたと感じています。有意義な時間でした。**

**(1) 料理領域検出部門  
最高精度賞 1位**

**(1) 料理領域検出部門  
最高精度賞 1位**

**チーム オリンパス・Eren**

## （1 料理領域検出部門）最高精度賞

### オープン

- ・YOLO物体認識フレームワークを利用
- ・全世界Deep Learningコミュニティの活発、技術のキャッチアップが重要

### データ

- ・End-to-Endな学習、再現性があり
- ・画像と正解ラベルさえ用意できれば、良い結果が得られる

### ノウハウ

- ・今回は3回の繰り返しFine-tuning学習を実施
- ・データの収集管理、製品化向け圧縮と高速化、法律面
- ・日本でのDeep Learning技術者間の交流をもっと

オリンパス・Eren チーム代表：李禎

**(1) 料理領域検出部門  
アイデア賞 1位**

**(1) 料理領域検出部門  
アイデア賞 1位**

**相崎 友保 (Izktmys)**

人工知能技術戦略会議主催  
AIチャレンジコンテスト（第1回）

(1)料理領域検出部門レポート

YOLOv2による料理領域検出

相崎 友保

# 一般物体検出アルゴリズムによる料理領域検出

## ■ Deep Learningによる一般物体検出アルゴリズムであるYOLOv2<sup>†</sup>を実装

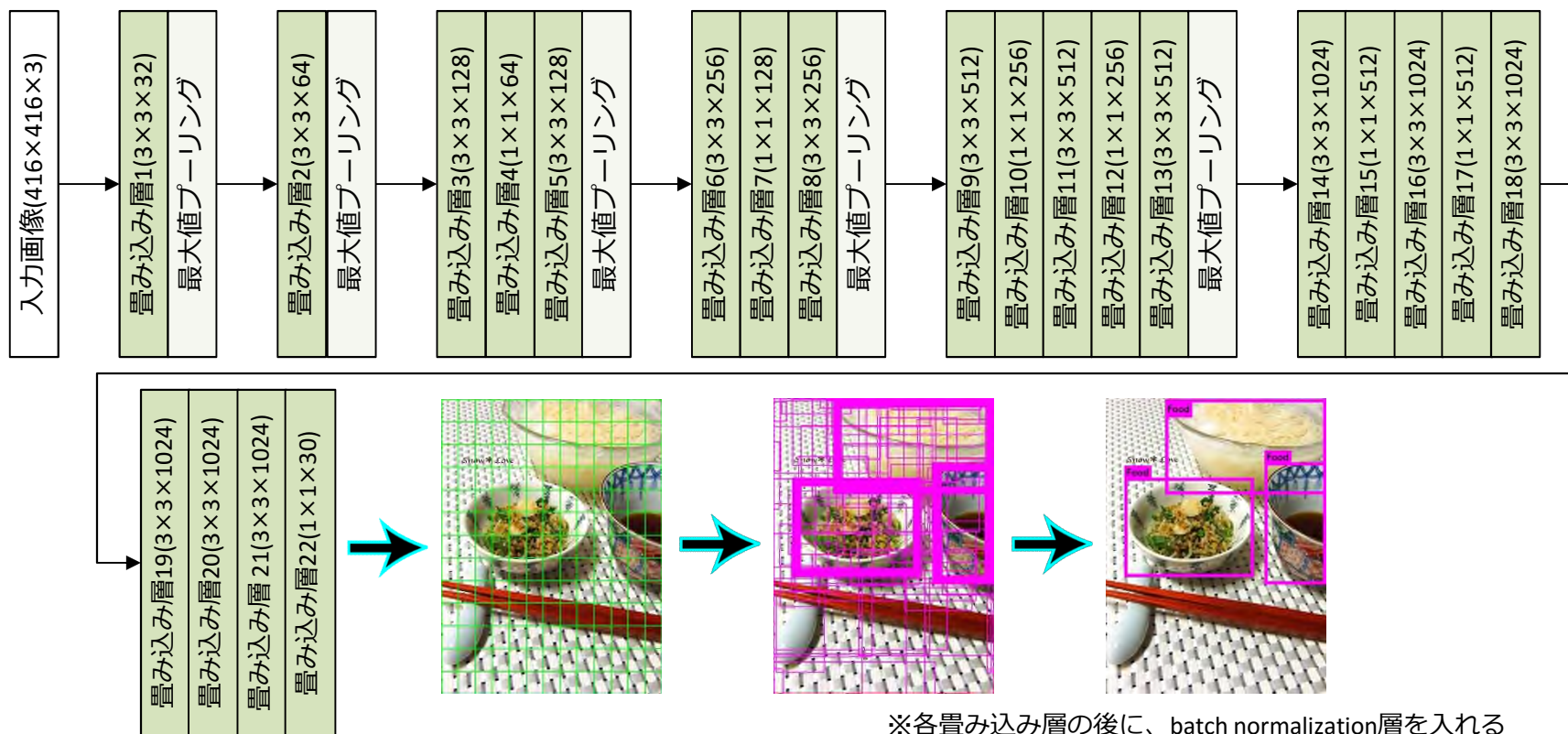
<sup>†</sup> : Joseph Redmon, Ali Farhadi, "YOLO9000: Better, Faster, Stronger" arXiv:1612.08242v1 [cs.CV] 25 Dec 2016

【特徴】ワンショット方式の物体検出CNNで、高速

【概要】ブロックに切り分け、ブロック毎にそのブロック内に中心がある

「矩形座標」と「矩形内に物体が存在する確率」、「物体のクラス確率」を算出

(フィルタ幅×フィルタ高さ×出力チャンネル数)



※各畳み込み層の後に、batch normalization層を入れる  
※各batch normalization層の後に、bias層を入れる  
※各bias層の後に、活性化関数leaky reluを使用



# データ利用方法の工夫点

## ■ 左右反転による訓練データの増量



オリジナル画像



左右反転画像

## ■ データクレンジング

- 料理画像以外を削除  
例) train\_263.jpg
- 画像サイズよりも料理領域サイズの方が大きいデータを削除  
例) train\_4678.jpg
  - 画像サイズ ; 280×273 < 料理領域サイズ ; 766×1023
- 料理領域サイズが0であるデータを削除  
例) train\_4234.jpg
  - 料理領域 (x:686, y:136, width:0, height:0)
- 料理領域の座標が負のデータを修正  
例) train\_63.jpg
  - 修正前 ; (x:-3, y:1, width:1001, height:1501)
  - 修正後 ; (x:0, y:1, width:998, height:1501)



train\_263.jpg

# モデリングの工夫点①

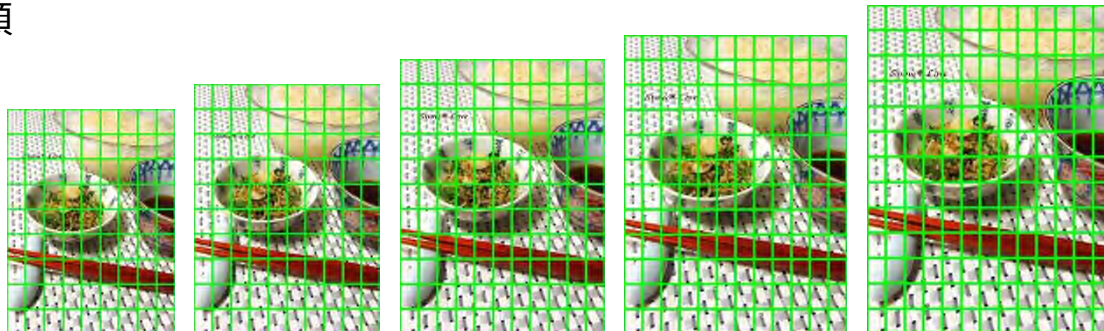
## ■ Dimension Cluster

- k-meansクラスタリングを用いて「よく使われるanchor boxのサイズ」を事前に決定する手法
- 13 × 13のブロック、 K = 5でk-meansを実行
  - width:4.30, height:3.81
  - width:7.48, height:9.25
  - width:9.34, height:4.50
  - width:12.51, height:9.24
  - width:12.69, height:12.48

## ■ Multi-Scale-Training

- YOLOv2はFully Convolutional Networks構造のため入力画像のサイズは可変
- 学習時に複数サイズの画像を交互に入力する事でロバスト性向上
  - 10イテレーション毎に画像サイズを変更
  - 画像サイズは5種類

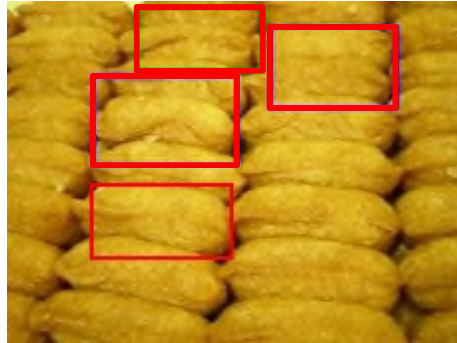
- 288 × 288
- 320 × 320
- 352 × 352
- 384 × 384
- 416 × 416



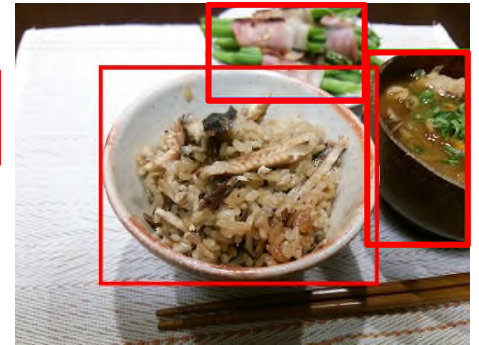
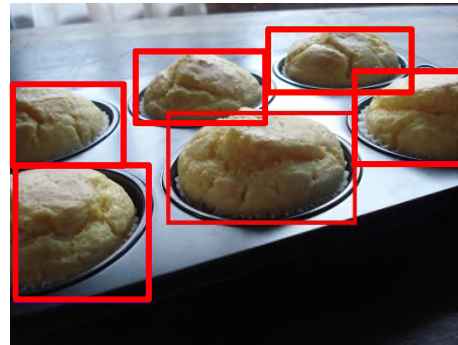
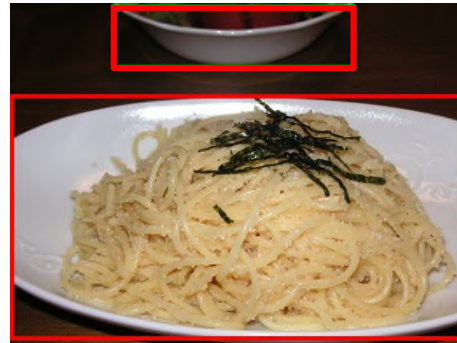
## モデリングの工夫点②

【課題】 単一画像に複数の料理領域がある場合、検出精度が低い事が判明

【対策】 単一画像に複数の料理領域がある訓練データのみで再学習を実施



# 料理領域檢出結果



**(2) 料理分類部門  
最高精度賞 1位**

**(2) 料理分類部門  
最高精度賞 1位**

**チーム ケ～ン**

# AIチャレンジコンテスト

チーム：ケ〜ン

料理分類部門：1位

料理領域検出部門：2位

東京大学 相澤・山崎研究室

伊神 大貴, 井上 直人, 小川 徹, 竹木 章人,  
松井 勇佑, 山崎俊彦, 相澤 清晴

# 料理分類部門

---

-25クラス

-10,000枚のラベル付き画像

-50,000枚のラベル無し画像



# 料理分類部門

---

-25クラス

-10,000枚のラベル付き画像

-50,000枚のラベル無し画像

→頑張っても使っても精度上がらない...

# 提案手法

---

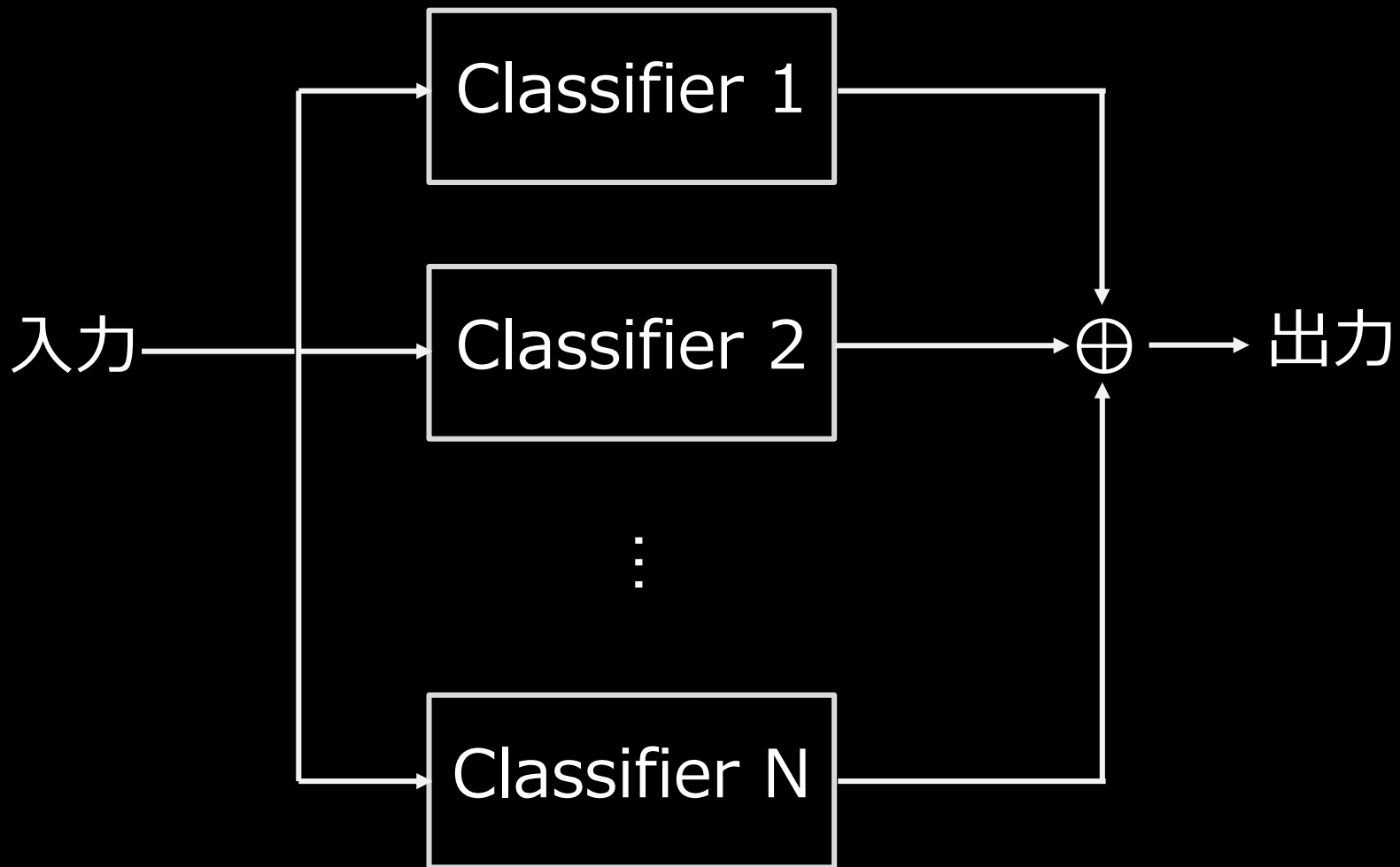
でいーぷらーにんぐ

10,000枚のラベル付き画像のみ使う

(GPU) Power is Justice!!

# アンサンブル学習

独立に識別器を学習し，予測結果を統合する



# Power is Justice.

Model	使用した個数	出力の重み
ResNet50-A	2	1/3
ResNet50-B	1	1/3
ResNet101	1	1
ResNet152	1	1
ResNet200	1	1
DenseNet161	5	1/5
DenseNet169	5	1/5
DenseNet201	5	1/5
PyramidNet300	5	1/5
PyramidNet450	5	1/5

31個のモデルのアンサンブル

# 結論

---

綺麗なデータセットは重要

# 結論

---

綺麗なデータセットは重要

(GPU) Power is Justice!!

# 結論

---

綺麗なデータセットは重要

(GPU) Power is Justice!!

(GPU) Power is Justice!!

**(2) 料理分類部門  
アイデア賞 1位**



**(2) 料理分類部門  
アイデア賞 1位**

**増田 知彰 (Tomo)**

# 第1回 AIチャレンジコンテスト 料理分類部門 アイデア賞

-ゼロからいかに深層学習を学ぶか-

増田 知彰

# 自己紹介

NTTコミュニケーションズ (※コンテストは個人参加)

増田 知彰 |  @tomo\_makes

オン:

IoT関連のBizDev

Product Management

オフ:

Weekend Maker/Programmer

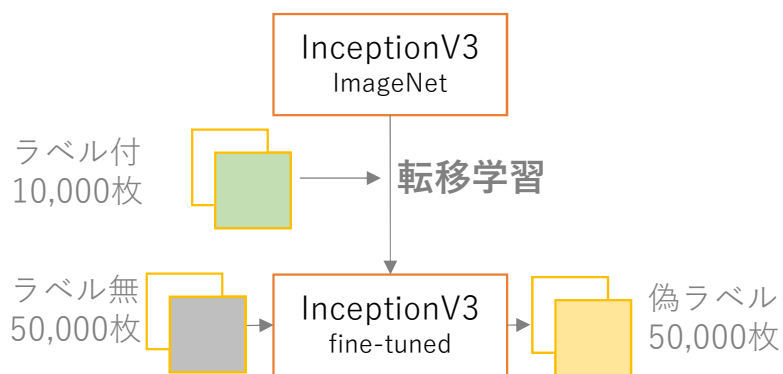
データ分析/機械学習

OSS活動(非コーディング主体)

# 最終モデルと結果 10<sup>th</sup>/138 (78.4%)

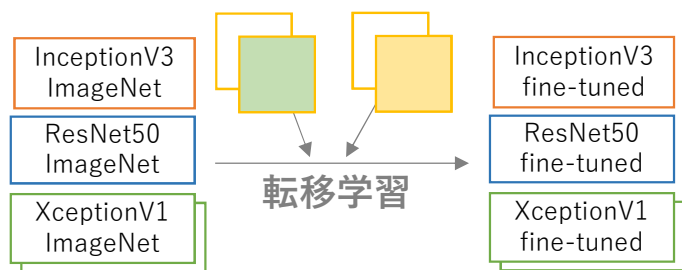
Keras/TensorFlowで、Pseudo-labelling + Knowledge distillation, Blending(6model出力線形回帰)を実施

## 1. Pseudo-labelling

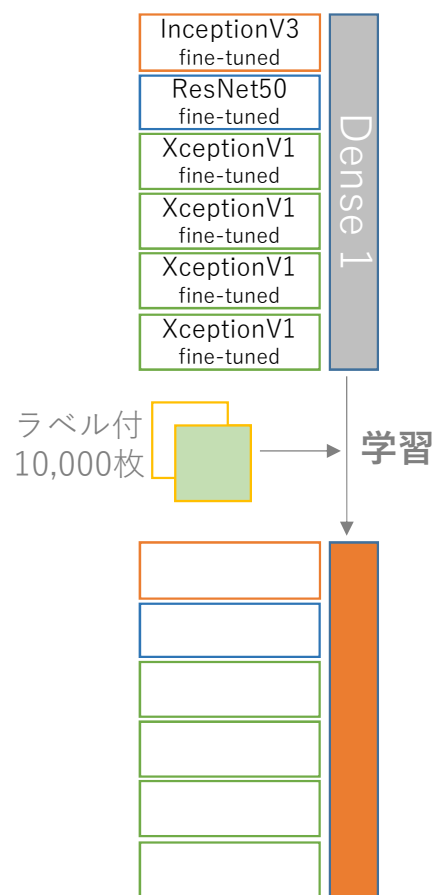


## 2. Knowledge distillation

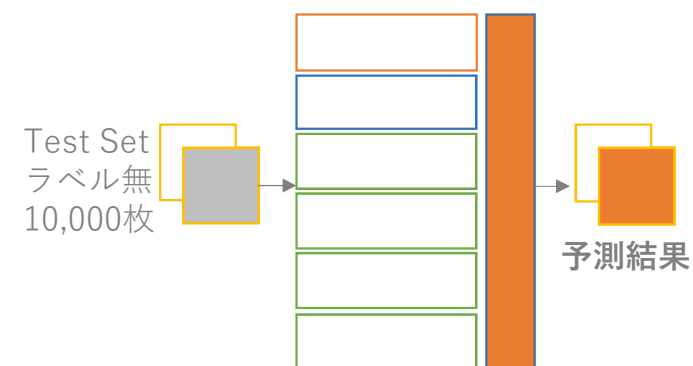
ラベル付、偽ラベルを4:1~3:1で混合  
バッチ学習し、6モデル生成



## 3. Blending 6 models



## 4. Predict



## Tips

- ラベル付9,000枚を学習/1,000枚を検証用
- 各学習時のReal-time Augmentation (Test-time Augmentationは効果なし)
- L2正則化、バッチ正則化使用
- SGD学習率固定 or スケジューリング
- Jupyter Notebookの活用
- Slackによる学習監視 など



でも、  
2ヶ月前は何ひとつ  
知りませんでした

# 精度 4%から78%にいたる2ヶ月

ゼロからいかに深層学習を学ぶか

学習意欲

(ハンズオンでできること/  
できないことを見極めたい)

+

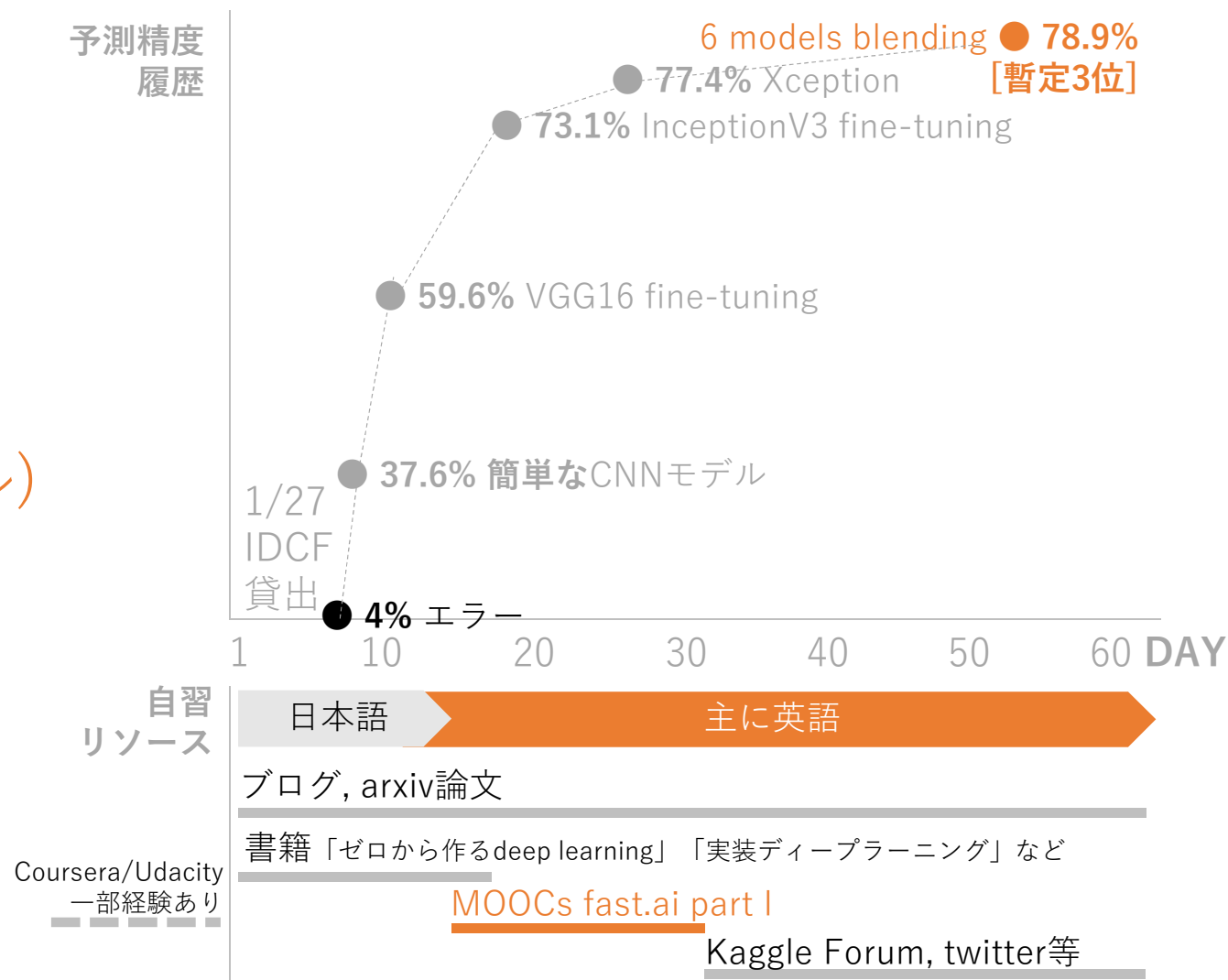
各段階に必要な自習リソース  
(自己課題に適用できるレベル)

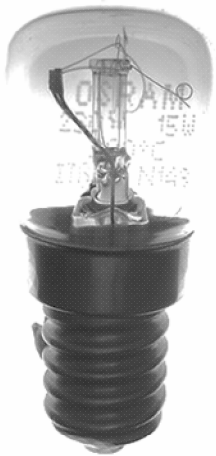
+

財布を気にせぬGPUリソース  
を使った実践

+

限られた時間、刻々と変わる  
ランキングの刺激





それならみんなできるのでは？

Repeatableにするために必要なこと

- **実務/中級レベルリソース**へのアクセス
- 実践を通じた学習機会



# 取り組み予定

2月の私と同じように学ぼうとした人が、最短距離で自分の課題に適用できるレベルへ進めるように。  
より日本で人工知能活用が熟し、かつ諸外国に遅れをとらないためにできること

- fast.ai Kaggle元CSO Jeremyの実践的Deep Learningコース. University of San Franciscoでの講義をMOOCs化



<http://course.fast.ai/>

- Part I (18hrs) 字幕**日本語訳完了**

- Part II (17hrs) 公開後翻訳予定 =実践的な実務/中級リソースが揃う

- コンテスト取組、学習経緯の記事化(権利関係が許す限り)

- 教育/交流の日本語コミュニティ化(構想段階)

活動への賛同、有形無形のご支援、または私がお役に立てる何かは

 @tomo\_makes までご連絡ください



# 人工知能技術戦略会議 議長賞

# 人工知能技術戦略会議 議長賞

チーム ケ～ン

人工知能技術戦略会議等主催  
第1回AIチャレンジコンテスト  
表彰式