

知的情報処理研究室

理工学部 應用情報工学科



Intelligent information precessing lab

iyatomi@hosei.ac.jp, <http://iyatomi-lab.info>

Deep Learning

研究テーマと目的

- ・コンピュータによる高度知的情報処理の実現
- ・人間には難しい問題に対して、支援を行うシステムの開発
- ・世の中を変える新しい技術の開発

構成人数

- ・教員 1名, 博士課程 2名
- ・大学院2年生 8名, 大学院1年生 7名
- ・学部4年生 12名 (留学生: ベトナム, 中国, エジプト)

- ・各種 画像処理
 - ・生命を模した新しい情報処理
 - ・ネットワーク技術
- 等を切り口として新しい価値の創造を目指す!

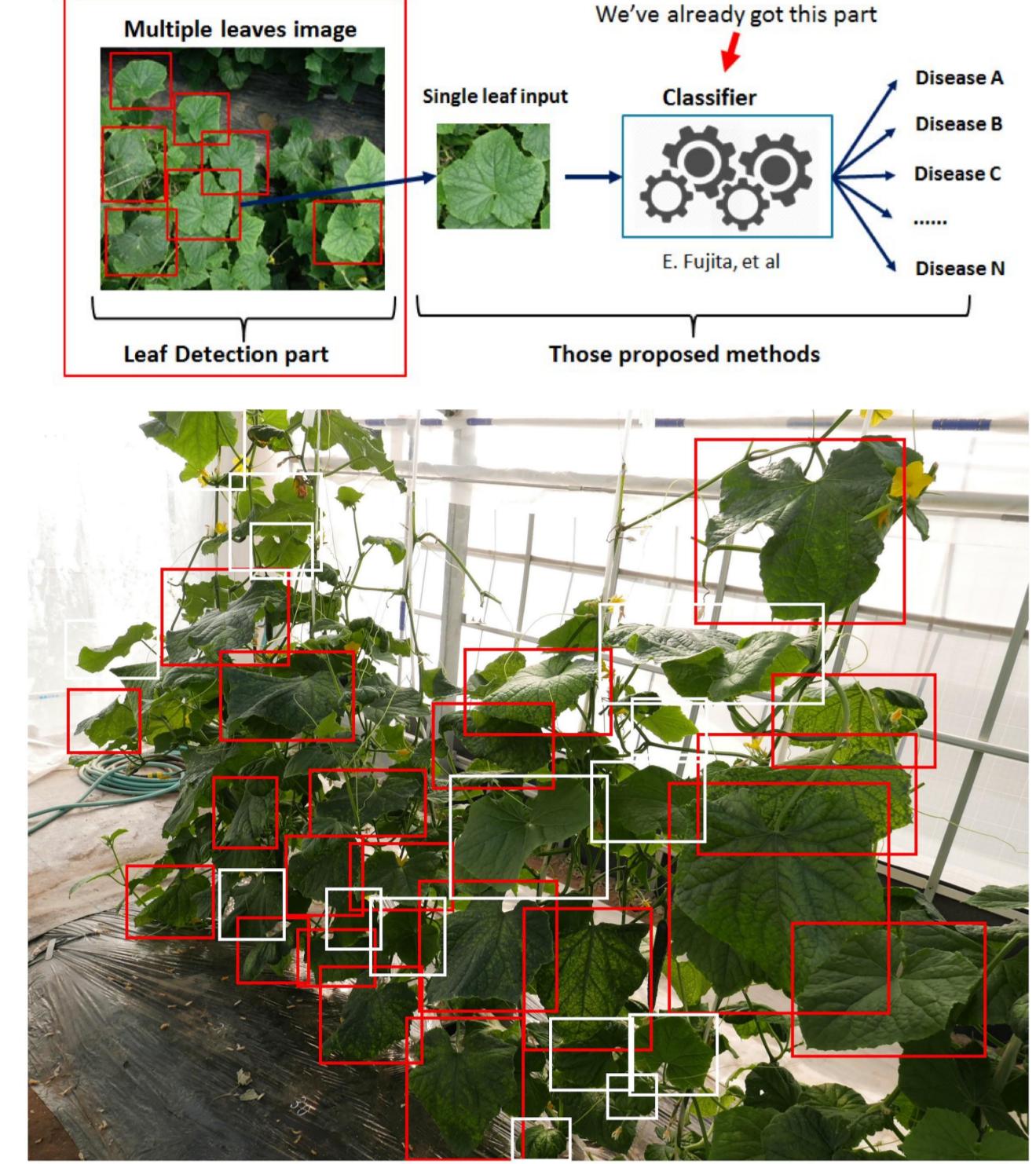
■ 深層学習を用いた植物病自動診断

24府県の農業試験場, 農研機構(国研)等と共同で植物病自動診断システムの構築

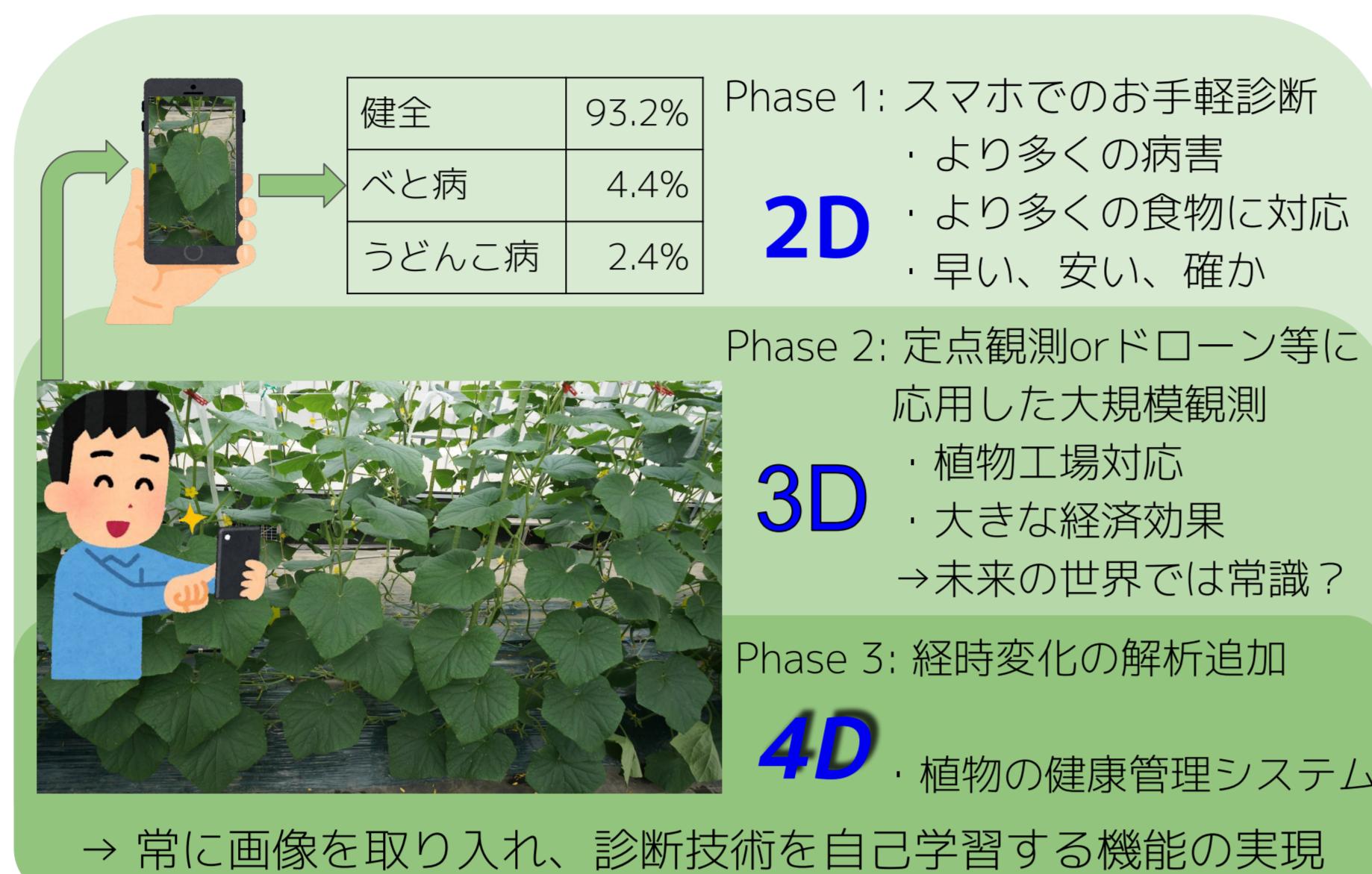
農研機構 NTT DATA

日本農業株式会社 Northern system service

深層学習を用いた高速な葉領域抽出において84%の精度を達成

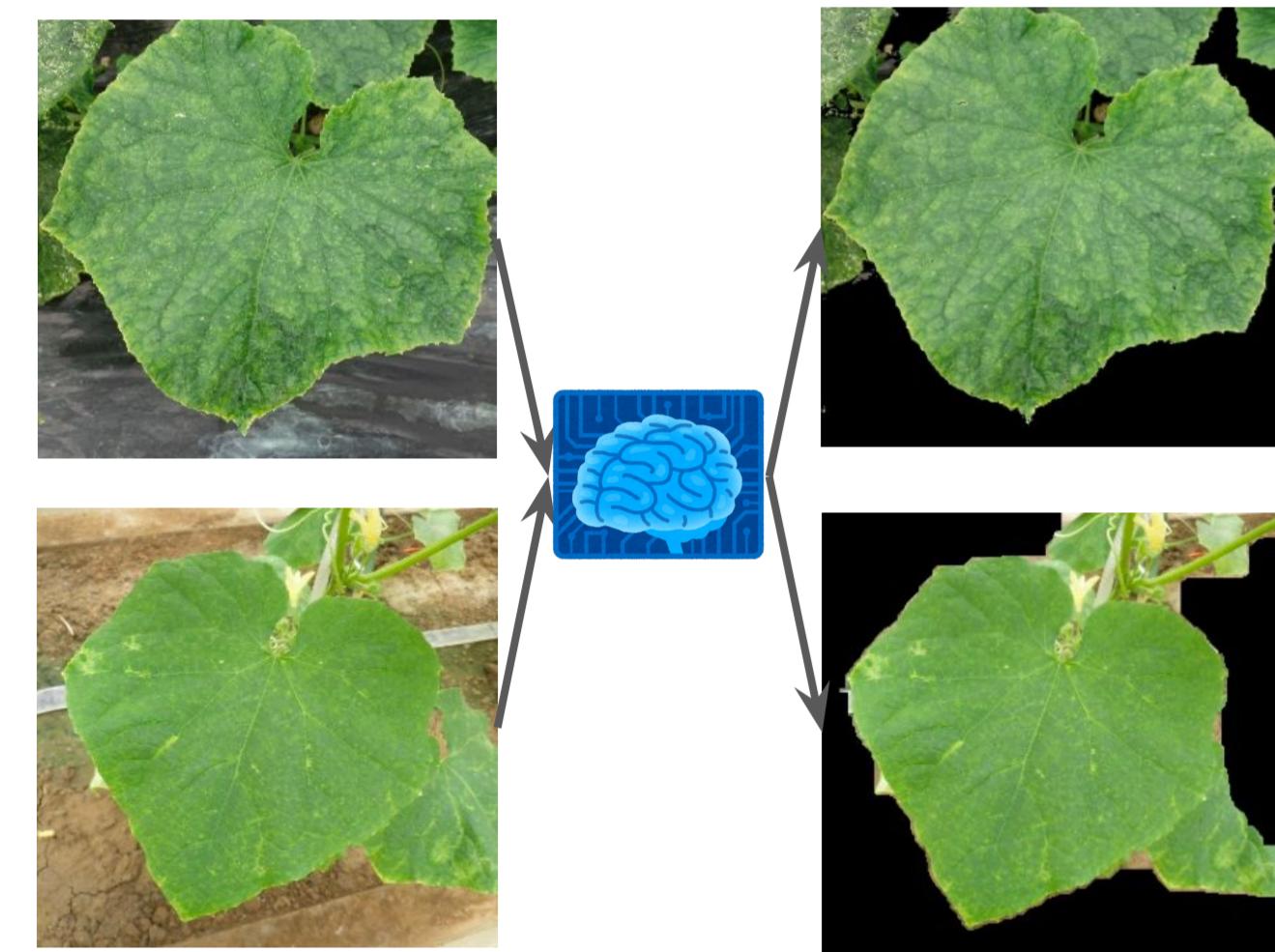


スマートフォンを用いたお手軽診断



農場に左右されない背景除去による頑健な識別器の学習

過学習を抑制し実環境における精度4%向上



複合感染したキュウリの自動診断 IEEE AIPR 2018

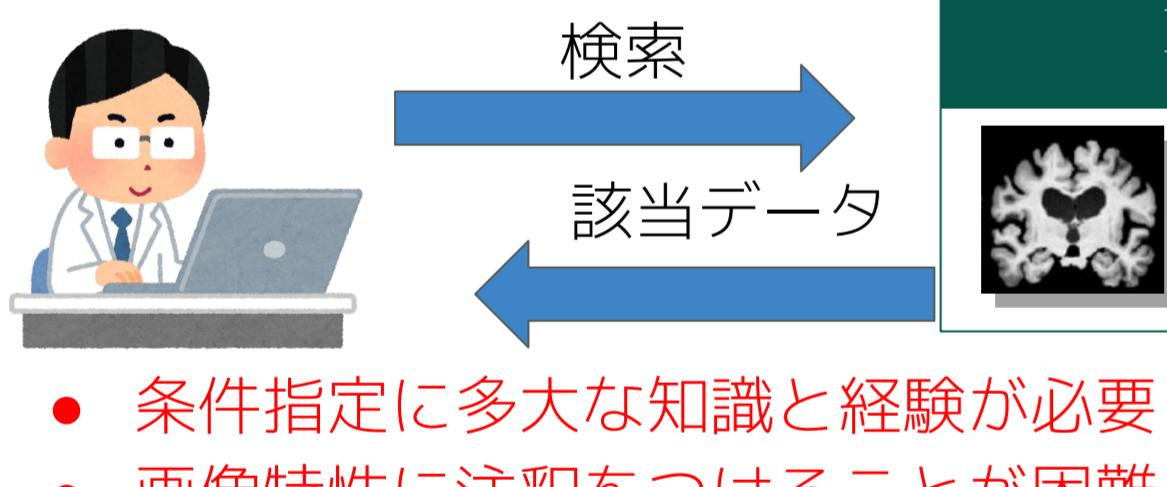


■ コンテンツベースの深層画像検索

"Deep" Content-Based Image Retrieval (CBIR) IEEE EMBC 2018

最先端の医療現場からの要望
世界的な専門医でも診断が難しい

Objective: 大規模なMRI画像群から、
過去の似た症例を検索したい!

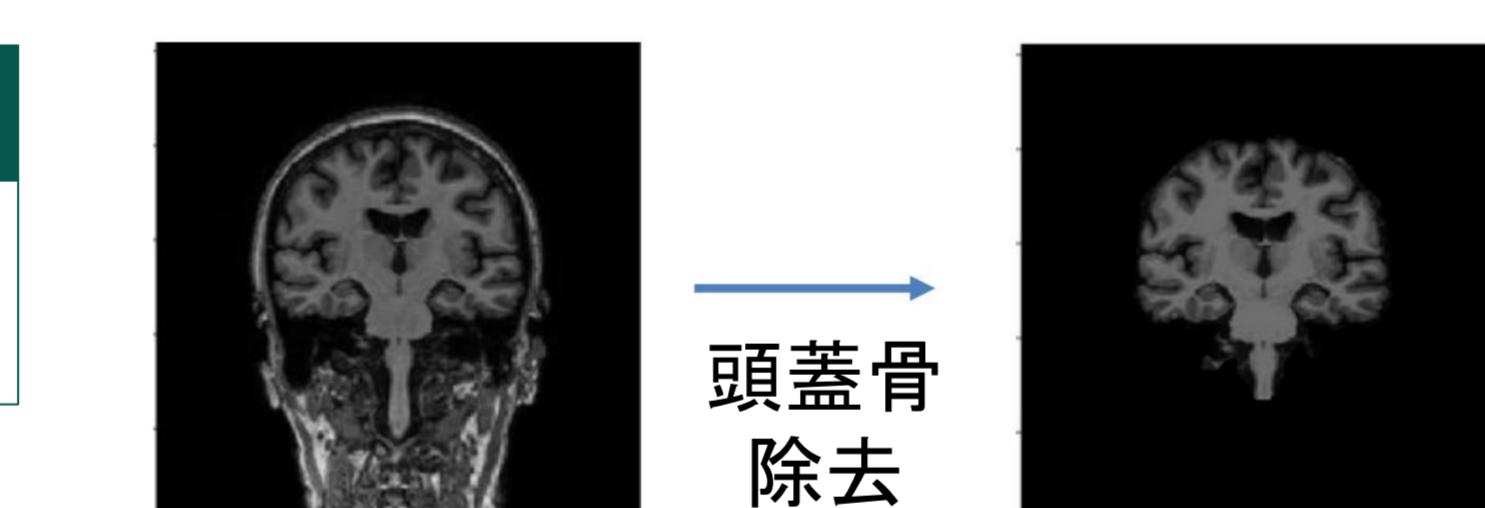


- ・条件指定に多大な知識と経験が必要
- ・画像特性に注釈をつけることが困難

敵対的生成ネットワークによるMRI脳画像の頭蓋骨除去

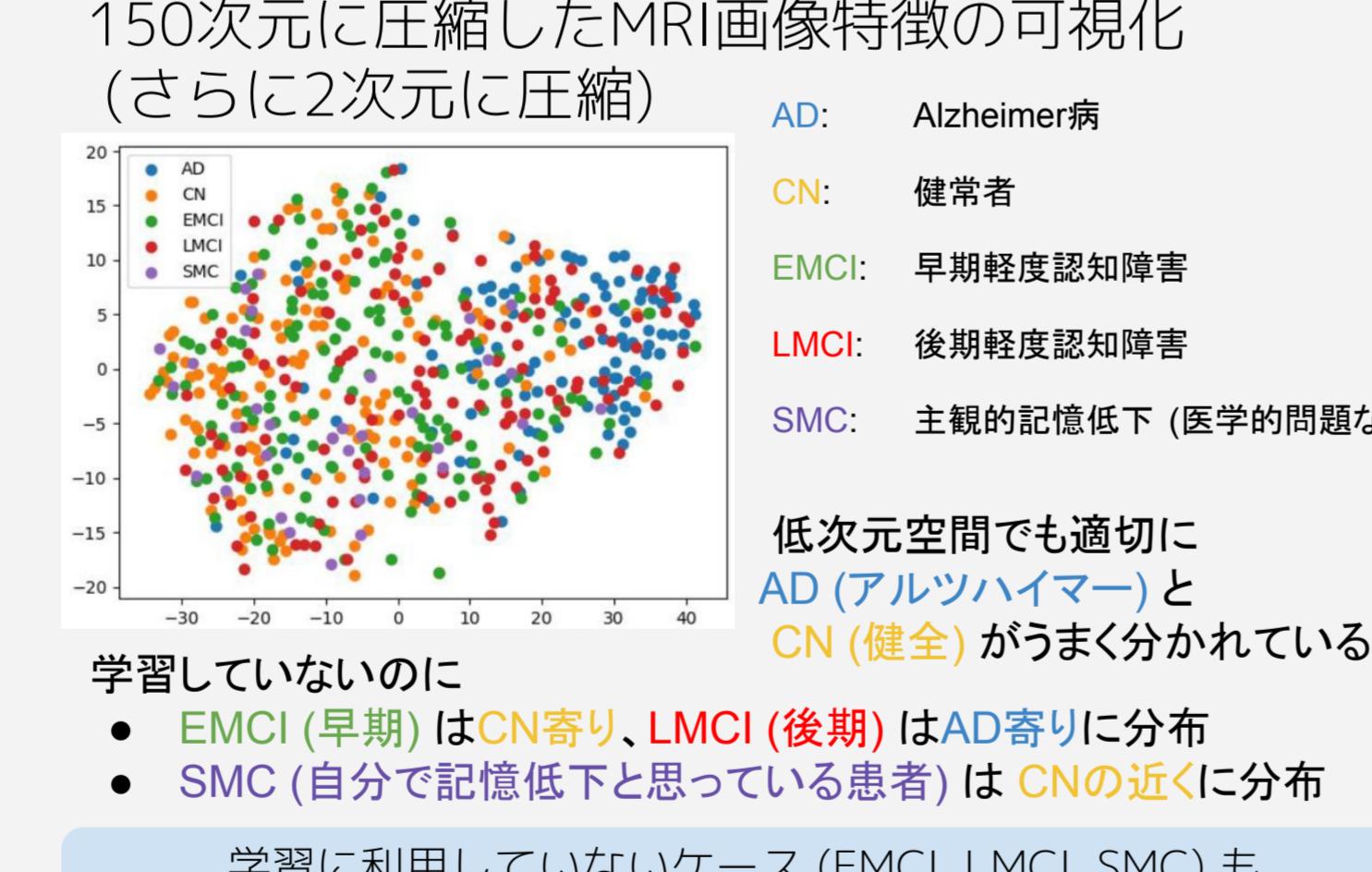
従来手法では除去精度や処理時間に課題
提案手法により F1スコア 平均97.28%

1症例当たりの生成時間は 約3秒 (従来の400倍高速)



3D-convolutional auto-encoders(3D-CAE)を用いた次元削減

150次元に圧縮したMRI画像特徴の可視化
(さらに2次元に圧縮)



診断に必要な情報を保持したまま たった150次元で表現!

150次元まで圧縮後の再構成画像: RMSE 約6%を実現

■ 深層学習を用いた独創的な自然言語処理

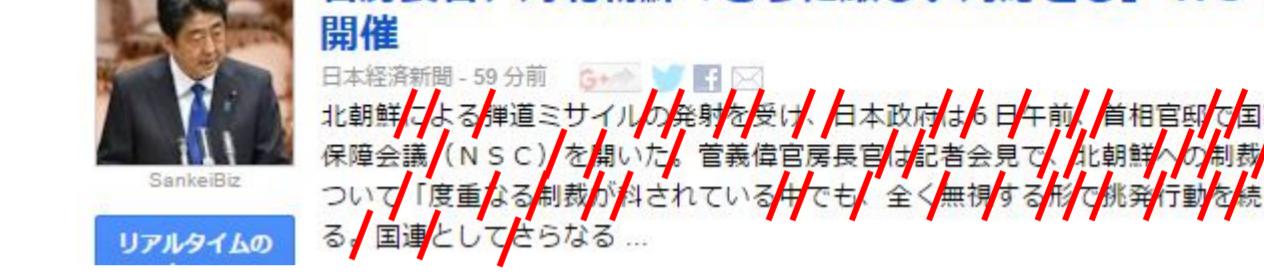
文字“画像”を用いた自然言語処理

従来(英語): 単語に分割 → 単語IDを元にベクトル化 → ベクトル表現からパターンを学習

日本語は使う文字種が英語と比べてとても多い

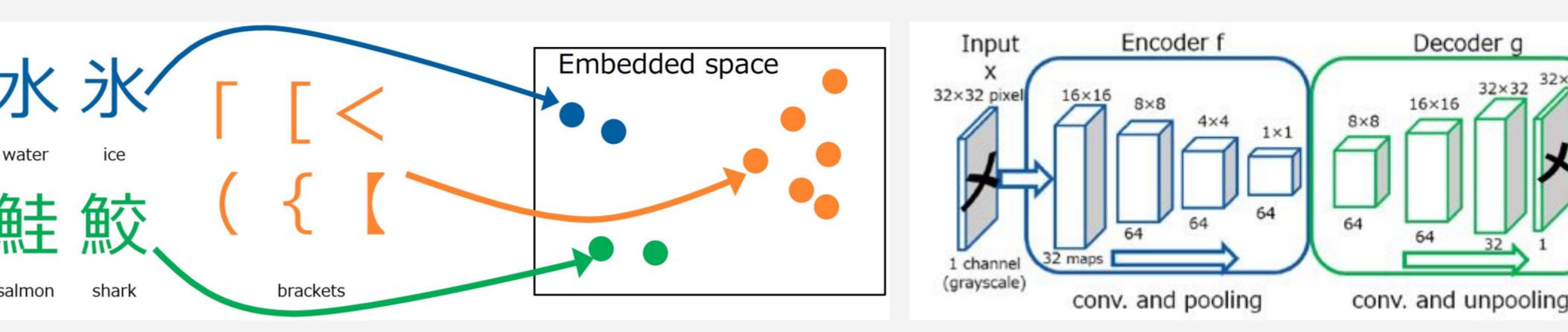


形態素解析を用いても文章を適切に分離するの困難



日本語の文字の表意性に着目・文字単位で処理することで単語分割不要な

文字の形状的特徴を利用した文書分類法を提案 IEEE BigData 2016, IEEE AIPR 2018



さらに発展させて、
文章の各文字を画像群と表現、
文書から識別までをend2endで
行うモデルの開発

- ・画像空間と特徴空間の両方で data augmentation可能
- ・画像処理で得られた知見を 文書処理に応用可能
- ・CNNベースなので 高速・並列化が容易

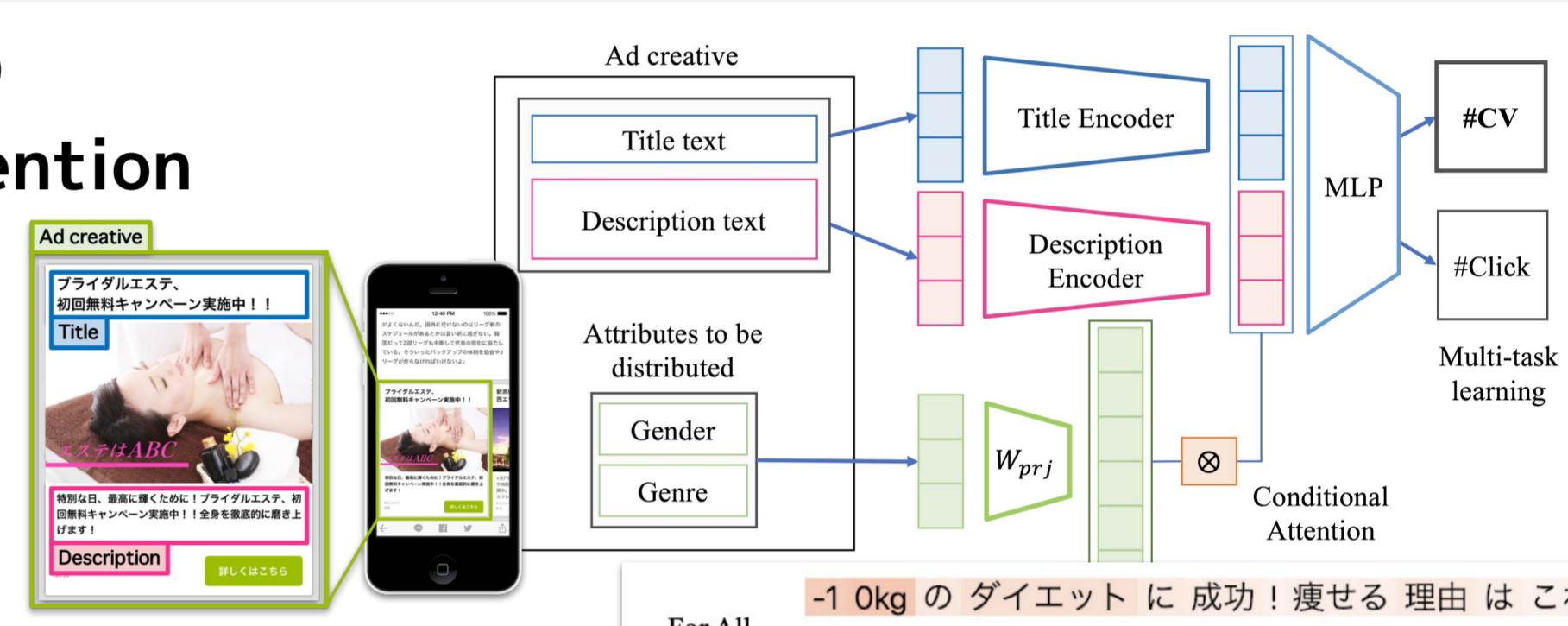
Wikipediaタイトルカテゴリ
分類 (206Kデータ、12クラス)
で世界最高記録を達成

広告クリエイティブ作成支援のためのマルチタスク学習とconditional attention

によるコンバージョン予測 KDD 2019
(データマイニングの最難関国際学会)

広告の効果予測と広告作成支援のための
新しいフレームワークの提案

- ・マルチタスク学習: 不均衡なコンバージョン
に対してクリックを同時に予測することで精度向上
- ・Conditional attention: 配信対象の属性を考慮した新しい注意機構
注意機構を用いた重要語の可視化をもとにしたクリエイティブ作成支援



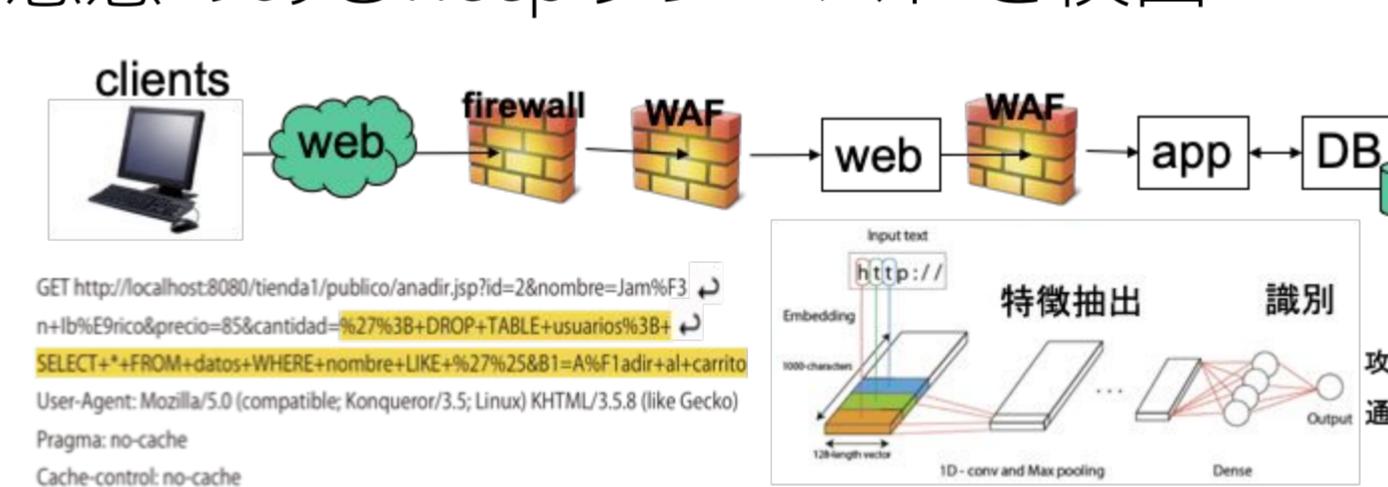
- For All: -1.0kg のダイエットに成功! 減せる理由はこれ
女子に人気の方法で効果を実感
- For Women: -1.0kg のダイエットに成功! 減せる理由はこれ
女子に人気の方法で効果を実感
- For Men: -1.0kg のダイエットに成功! 減せる理由はこれ
女子に人気の方法で効果を実感

■ 深層学習 x セキュリティ

深層学習を用いた Web Application Firewall (WAF) の作成 IEEE CSPA 2018

従来のテンプレートを用いたblacklist方式、垂墨攻撃を防げない

悪意のあるhttpリクエストを検出



HTTP DATASET CSIC 2010 Dataset (normal 36k + attack 25k = 61k)

Detection accuracy: 98.8 % @ 2.35 ms (10 fold-cross valid)

同じデータセットこれまで報告されている最善の結果: 82.0 % with Naive Bayes

深層学習を用いた悪意のあるソフトウェア (malware) の解析 IEEE CSPA 2018

Malware解析: 新種や垂墨の増加が著しく解析が困難に

表層解析: ファイルの特徴を解析

動的解析: 実際に稼働させて動作を解析

静的解析: 機械語や命令を解析

× malware特徴の定義が困難

Character-level CNNを利用

malwareの隠れた特徴を

学習より獲得

- ・展開不要 (高速)
- ・世界最高レベルの識別精度

Android APKファイル (圧縮ファイル) そのものの
たった1024byteの解析のみで 93 ~ 96%の識別率を達成

■ 悪性黒色腫 自動診断システム

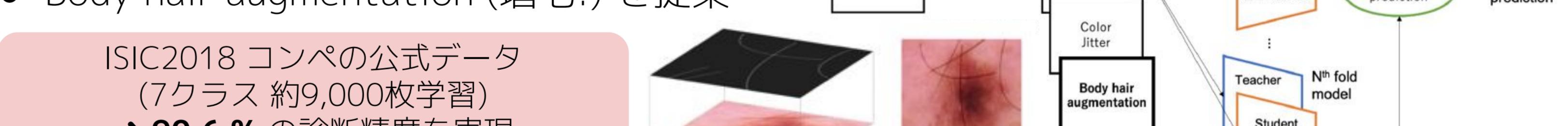
早期発見が大切だが診断が難しいメラノーマ診断支援

(医師の診断精度: 75-84%)

機械学習で最先端の

- ・SE-ResNet101 (101層CNN) をベースに
- ・Mean Teachers (半教師あり学習)
- ・Body hair augmentation (増毛!) を提案

ISIC2018 コンペの公式データ
(7クラス 約9,000枚学習)
→ 90.6 % の診断精度を実現



現時点までの対外的成果 (2019年8月現在)

- ・原著論文: 49件 (うち英文誌33件)
- ・国際会議論文: 69件
- ・国内研究会などでの発表: 100件+

現在の外部研究資金

- ・文部科学省 科学研究費補助金 基盤C (平成29-31年)
○ 深層学習と次元削減技術を用いた環境変化に
口パストな植物病害自動診断システムの開発

- ・農林水産省 受託プロジェクト
○ AIを活用した病害虫診断技術の開発 (平成29-33年)
+ 総務省 add on project PRISM (平成31-33年)

これまで他に

- ・科研費6件, JST (科学技術振興機構) の研究助成6件, 企業からの受託・寄附研究 数件

共同研究先

- ・Johns Hopkins University, The center of Imaging Science, Radiological Science
- ・University of Central Arkansas., Dept of Computer Science
- ・Loughborough Univ., Dept of Computer Science

- ・東京女子医科大学医学部 皮膚科学教室

- ・東海大学医学部 放射線科学教室

- ・埼玉県農林総合研究センター (含む 計24府県の農業試験場)

- ・国立研究開発法人 農研機構 他企業など



日本農業新聞2018年4月27日

■ 近年の受賞と研究室の対外的な活動

奨励賞 / YANS2019 (2019年8月)

- ・“解釈性向上のための注意機構と損失勾配に対する関連損失の導入”, 北田 俊輔, 瀧富 仁

学生奨励賞 (x2) / 情報処理学会 (2019年3月)

- ・“脳MRI画像における3次元オートエンコーダーによる特微量抽出の中間表現の改善”, 音賀 優楓, 藤山 眞梧, 新井 颯人, 茶山 祐亮, 瀧富 仁, 大石 健一

- ・“頑健な皮膚腫瘍診断支援のためのbody hair augmentation”, 北田 俊輔, 瀧富 仁

Best Paper Award / ISCLA (2018年7月)

- ・“An end-to-end practical plant disease diagnosis system for wide-angle cucumber images”, Q. H. Cap, K. Suwa, E. Fujita, S. Kagiwada, H. Uga and H. Iyatomi, 2018 International Symposium on Computational Intelligence and Application. (ISCLA2018)

Best Paper Award / IEEE CSPA (2018年3月)

- ・“Web Application Firewall using Character-level Convolutional Neural Network”, Michiaki Ito and Hitoshi Iyatomi, 14th IEEE International Colloquium on Signal Processing and its Application. (IEEE CSPA2018)

学生奨励賞 (x2) / 情報処理学会 (2017年3月)

- ・“転移学習を用いた植物病害自動診断システムの試み”, 藤田 恵梨香, 川崎 雄介, 鍵和田 聰, 宇賀 博之, 瀧富 仁
- ・“教師なし学習を導入した深層距離学習法”, 松岡 佑磨, 島田 大樹, 瀧富 仁

知的情報処理研究室

理工学部 應用情報工学科



Intelligent information precessing lab

iyatomi@hosei.ac.jp, <http://iyatomi-lab.info>

Deep Learning

研究テーマと目的

- ・コンピュータによる高度知的情報処理の実現
- ・人間には難しい問題に対して、支援を行うシステムの開発
- ・世の中を変える新しい技術の開発

構成人数

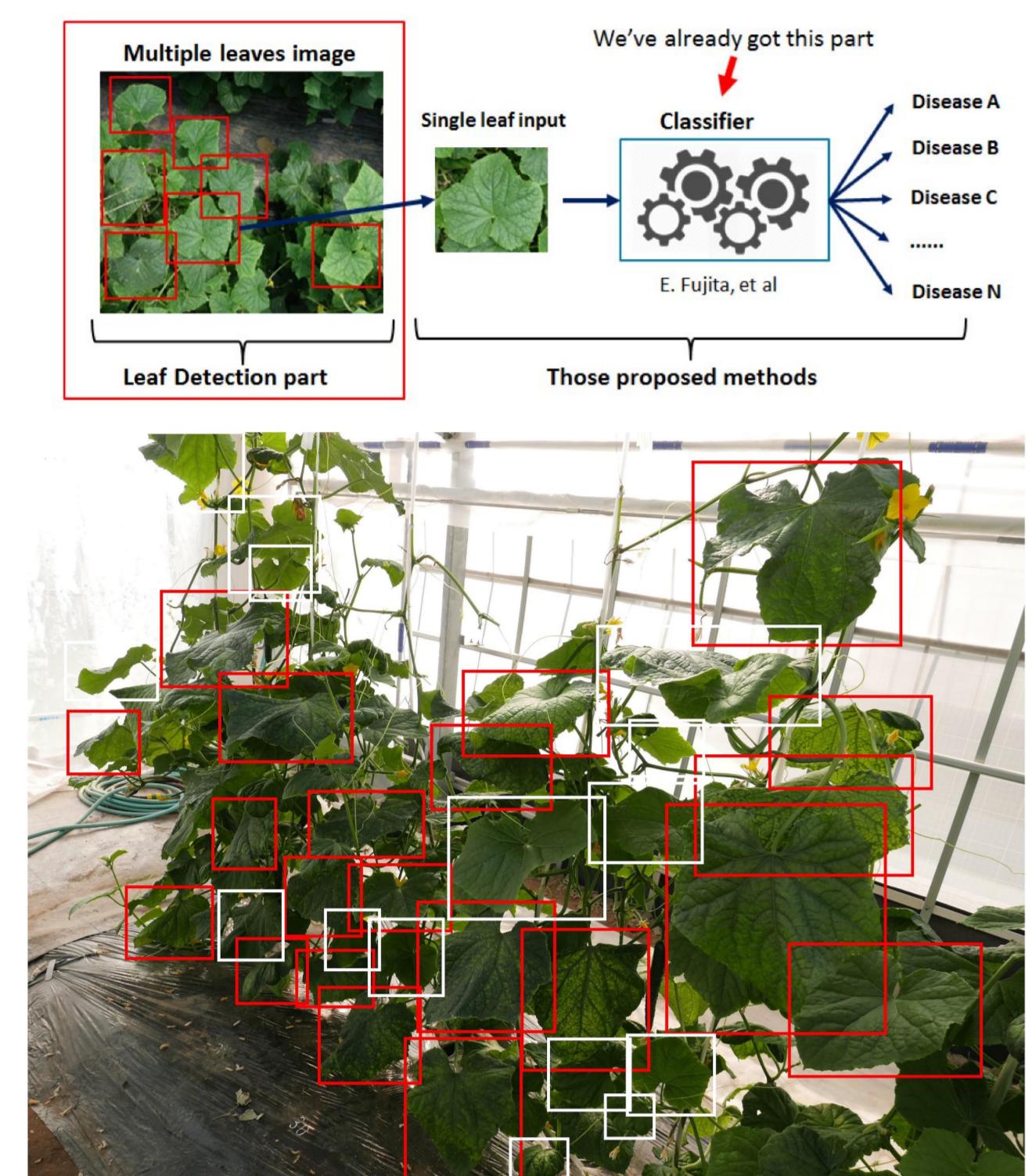
- ・教員 1名, 博士課程 2名
- ・大学院2年生 8名, 大学院1年生 7名
- ・学部4年生 12名 (留学生: ベトナム, 中国, エジプト)

- ・各種 画像処理
 - ・生命を模した新しい情報処理
 - ・ネットワーク技術
- 等を切り口として新しい価値の創造を目指す!

■ 深層学習を用いた植物病自動診断

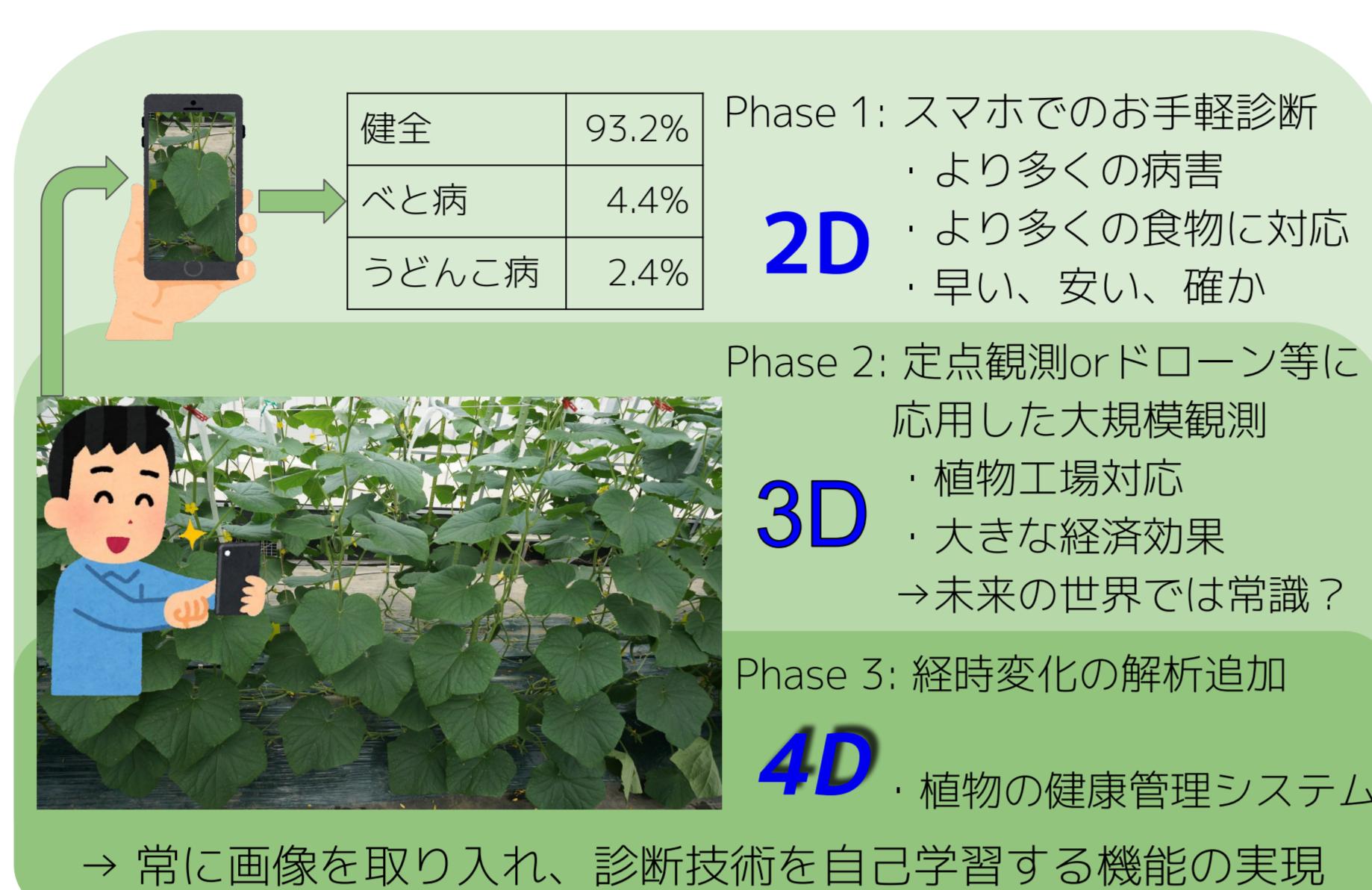
24府県の農業試験場, 農研機構(国研)等と共同で
植物病自動診断システムの構築

深層学習を用いた高速な葉領域抽出において84%の精度を達成

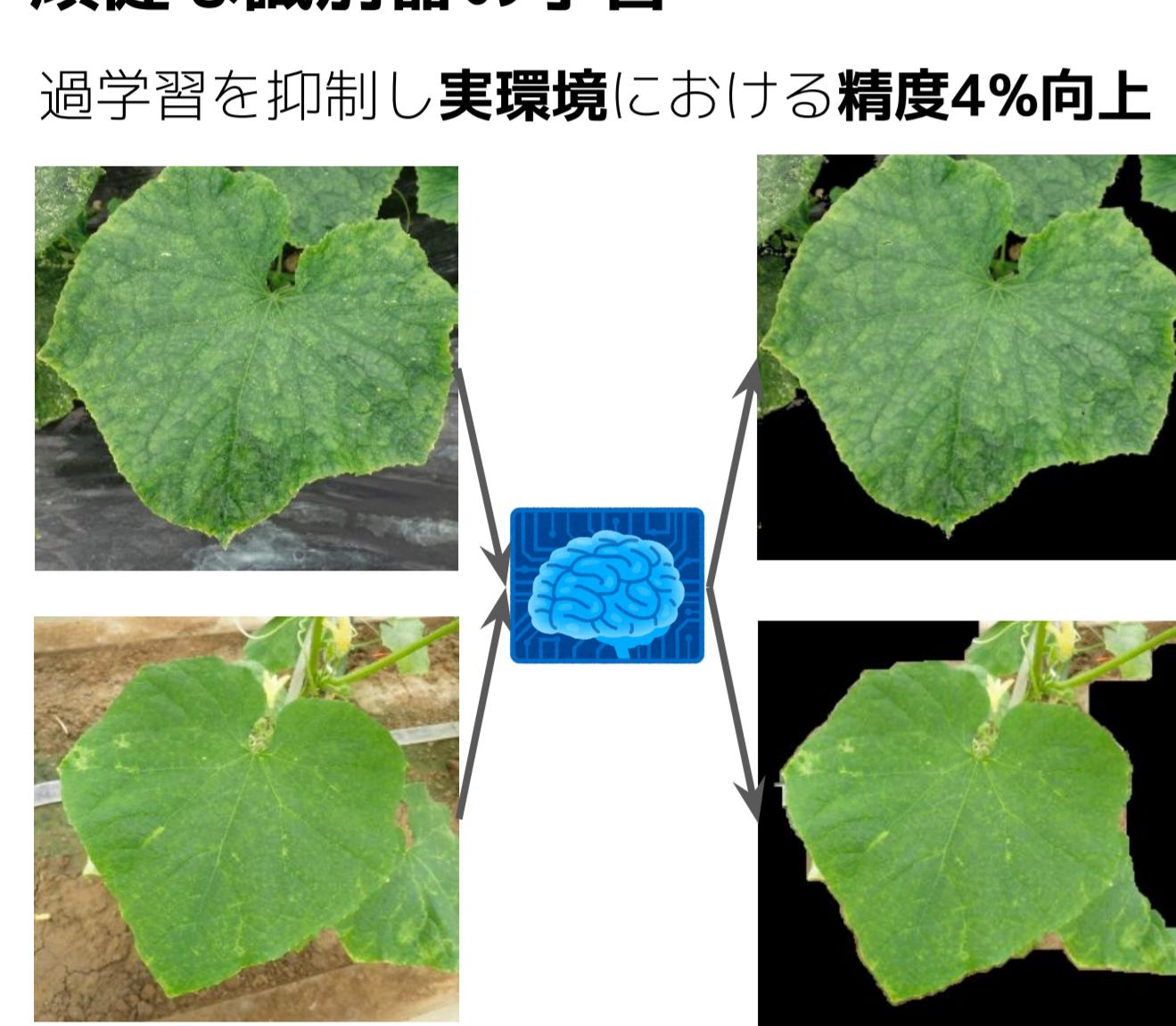


赤:トマト
黄:イチゴ
緑:キュウリ
紫:ナス

スマートフォンを用いたお手軽診断



農場に左右されない背景除去による
頑健な識別器の学習



複合感染したキュウリの自動診断 IEEE AIPR 2018

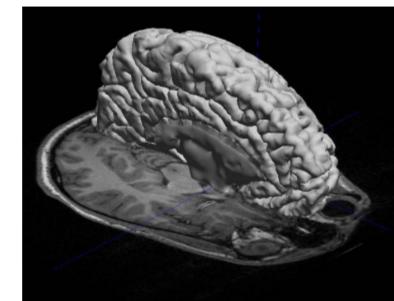
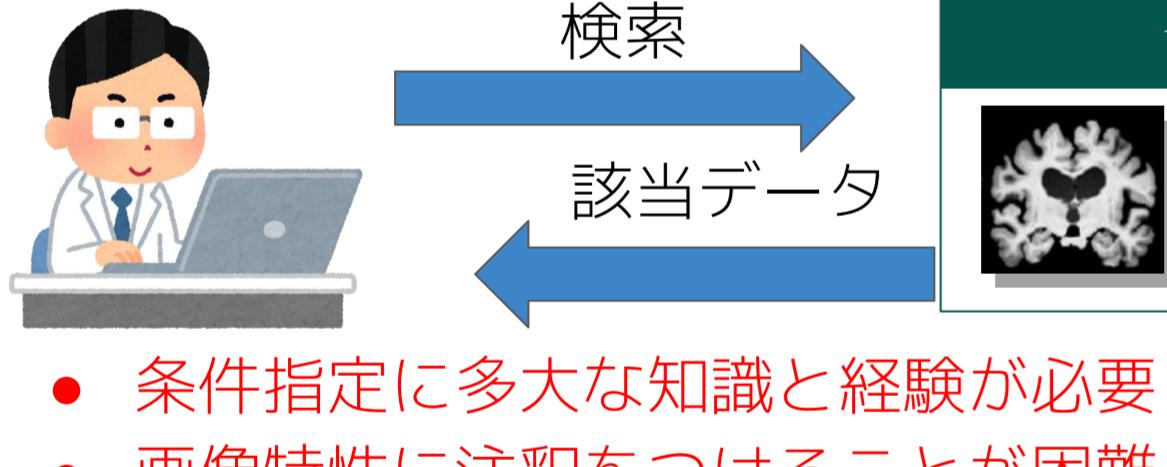


■ コンテンツベースの深層画像検索

"Deep" Content-Based Image Retrieval (CBIR) IEEE EMBC 2018

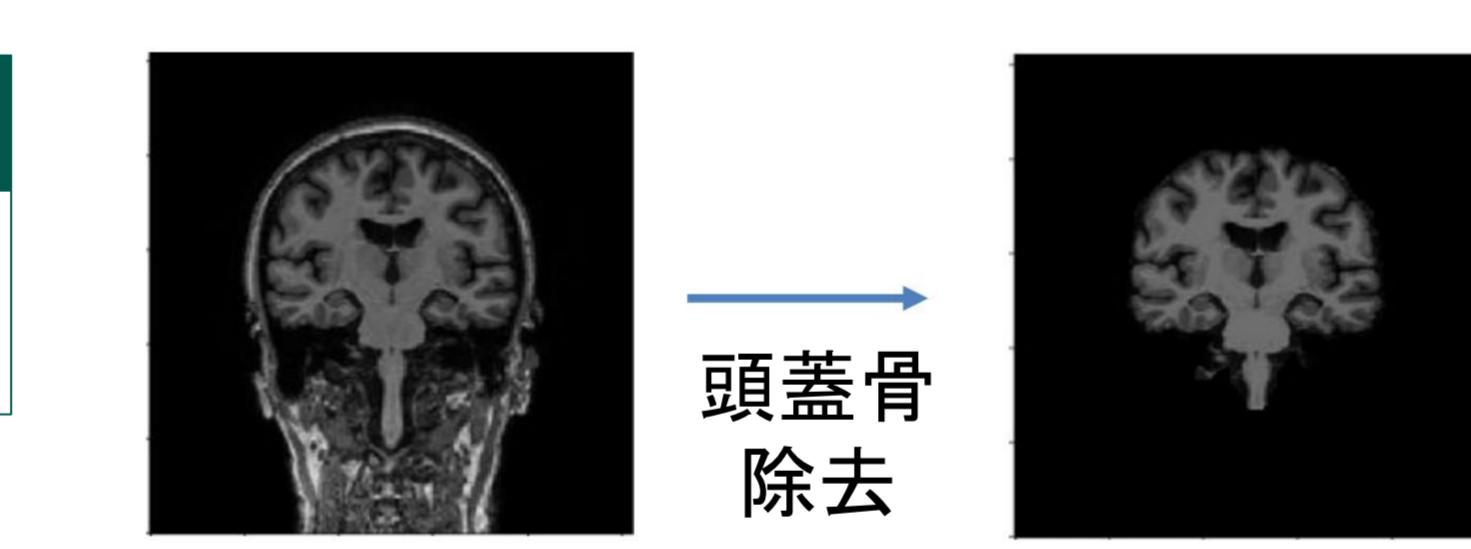
最先端の医療現場からの要望
世界的な専門医でも診断が難しい

Objective: 大規模なMRI画像群から、
過去の似た症例を検索したい!



敵対的生成ネットワークによる MRI脳画像の頭蓋骨除去

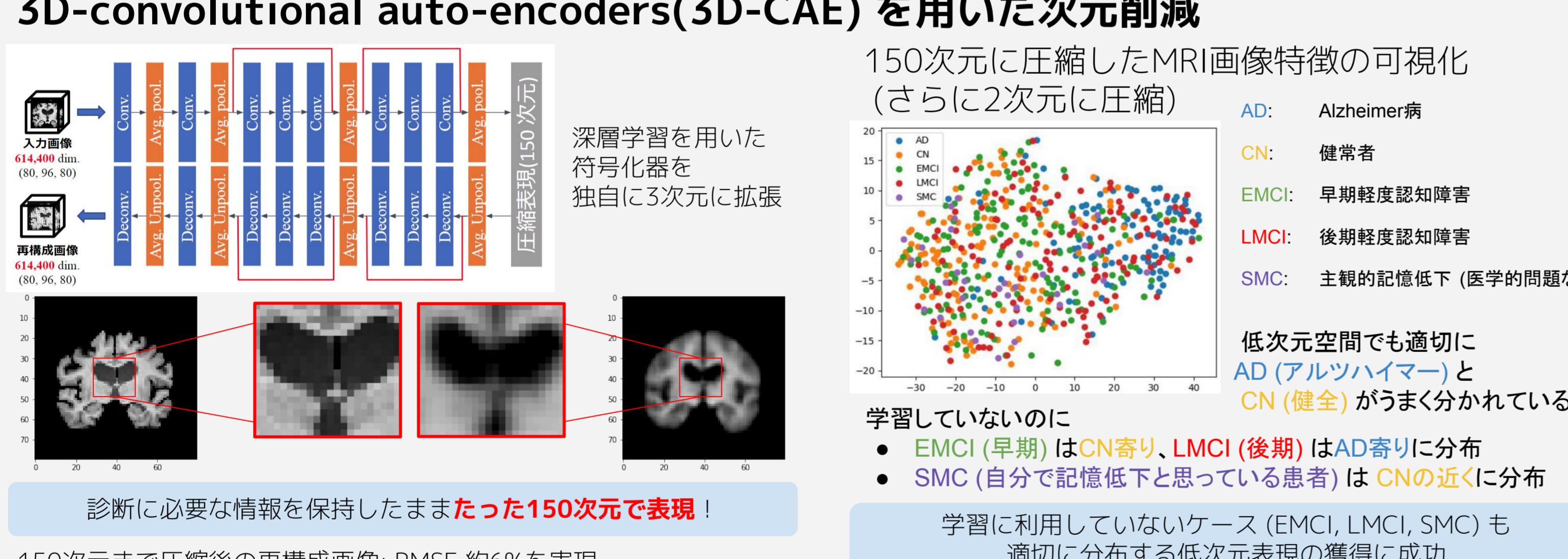
従来手法では除去精度や処理時間に課題
提案手法により F1スコア 平均97.28%
1症例当たりの生成時間は 約3秒 (従来の400倍高速)



頭蓋骨
除去

• 条件指定に多大な知識と経験が必要

• 画像特性に注釈をつけることが困難



■ 近年の受賞と研究室の対外的な活動

奨励賞 / YANS2019 (2019年 8月)

• "解釈性向上のための注意機構と損失勾配に対する関連損失の導入", 北田 俊輔, 研究員 仁

学生奨励賞 (x2) / 情報処理学会 (2019年 3月)

• "脳MRI画像における3次元オートエンコーダーによる特微量抽出の中間表現の改善", 音賀 優楓, 藤山 真悟, 新井 颯人, 茶山 祐亮, 研究員 仁, 大石 健一
• "頑健な皮膚腫瘍診断支援のためのbody hair augmentation", 北田 俊輔, 研究員 仁

Best Paper Award / ISCIA (2018年 7月)

• "An end-to-end practical plant disease diagnosis system for wide-angle cucumber images", Q. H. Cap, K. Suwa, E. Fujita, S. Kagiwada, H. Uga and H. Iyatomi, 2018 International Symposium on Computational Intelligence and Application. (ISCIA2018)

Best Paper Award / IEEE CSPA (2018年 3月)

• "Web Application Firewall using Character-level Convolutional Neural Network", Michiaki Ito and Hitoshi Iyatomi, 14th IEEE International Colloquium on Signal Processing and its Application. (IEEE CSPA2018)

学生奨励賞 (x2) / 情報処理学会 (2017年 3月)

• "転移学習を用いた植物病害自動診断システムの試み", 藤田 恵梨香, 川崎 雄介, 鍵和田 聰, 宇賀 博之, 研究員 仁

• "教師なし学習を導入した深層距離学習法", 松岡 佑磨, 島田 大樹, 研究員 仁

■ 深層学習を用いた独創的な自然言語処理

文字“画像”を用いた自然言語処理

従来(英語): 単語に分割 → 単語IDを元にベクトル化 → ベクトル表現からパターンを学習

日本語は使う文字種が英語と比べてとても多い

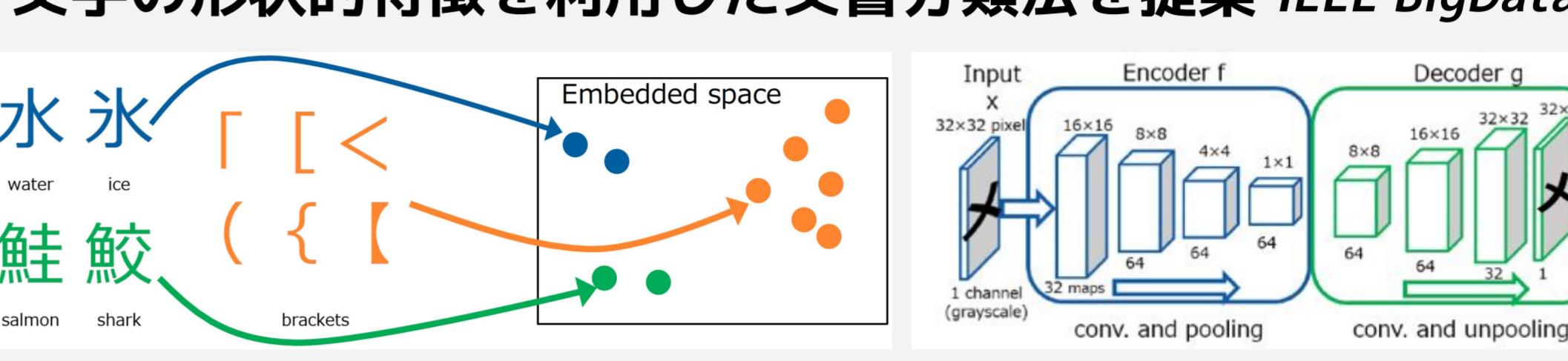
形態素解析を用いても文章を適切に分離するの困難

官房長官、対北朝鮮「さりに厳しい対応とする」 NSC 開催

北朝鮮は、北朝鮮の公文書の翻訳で、日本語の各単語をそのまま

日本語の文字の表意性に着目・文字単位で処理することで単語分割不要な

文字の形状的特徴を利用した文書分類法を提案 IEEE BigData 2016, IEEE AIPR 2018



似た形状を持つ文字を似たベクトルに写像

さらに単語分割が不要なdata augmentation

"wildcard training" を提案 (汎化性能向上!)

メロスは激怒した。→ メロス * 激 * した。

文章を学習する際、ところどころ文字をわざとnullに置き換え

さらに発展させて、
文章の各文字を画像群と表現、
文書から識別までをend2endで
行うモデルの開発

- ・画像空間と特徴空間の両方で data augmentation可能
- ・画像処理で得られた知見を 文書処理に応用可能
- ・CNNベースなので 高速・並列化が容易

Wikipediaタイトルカテゴリ
分類 (206Kデータ、12クラス)
で世界最高記録を達成

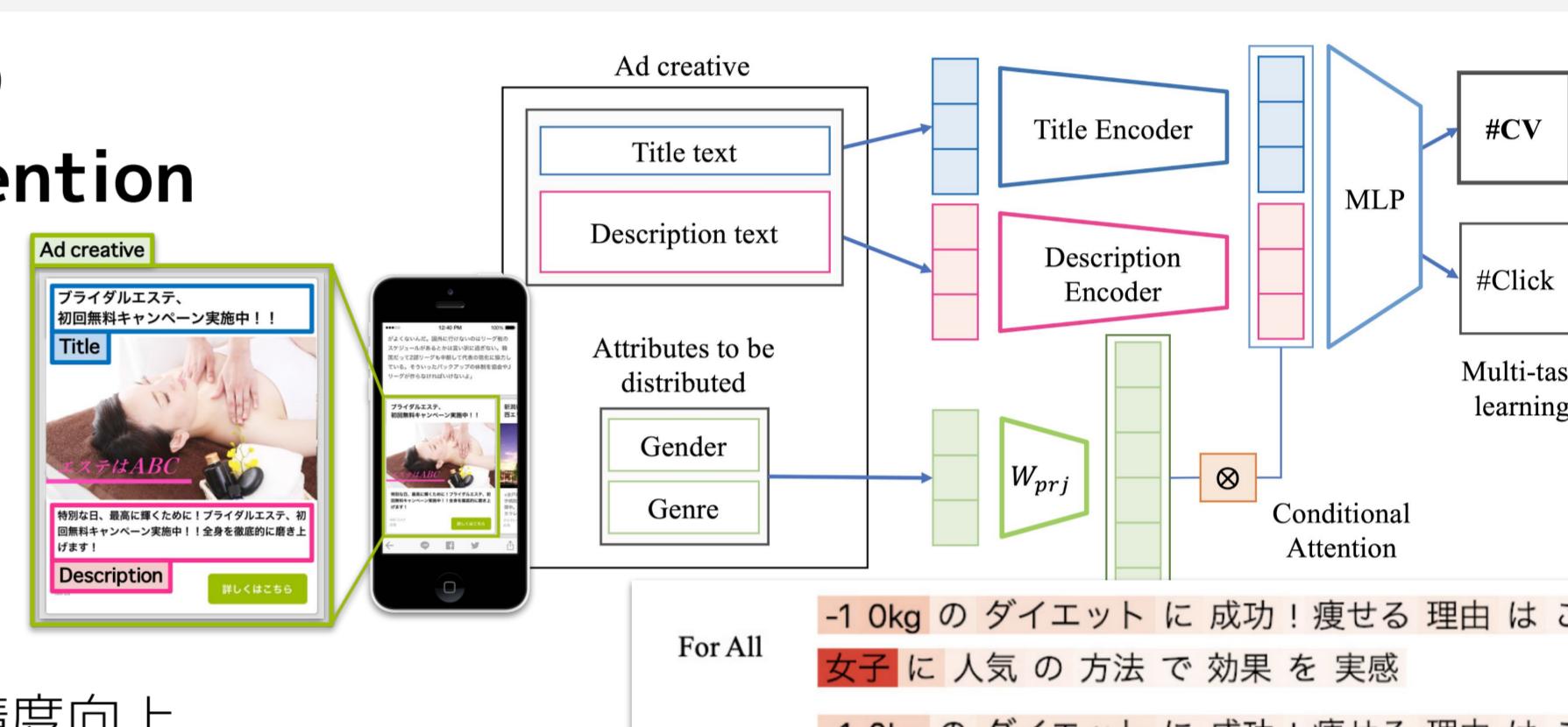
広告クリエイティブ作成支援のための マルチタスク学習とconditional attention

によるコンバージョン予測 KDD 2019

(データマイニングの最難関国際学会)

広告の効果予測と広告作成支援のための
新しいフレームワークの提案

- ・ **マルチタスク学習:** 不均衡なコンバージョン
に対してクリックを同時に予測することで精度向上
- ・ **Conditional attention:** 配信対象の属性を考慮した新しい注意機構
注意機構を用いた重要語の可視化をもとにしたクリエイティブ作成支援



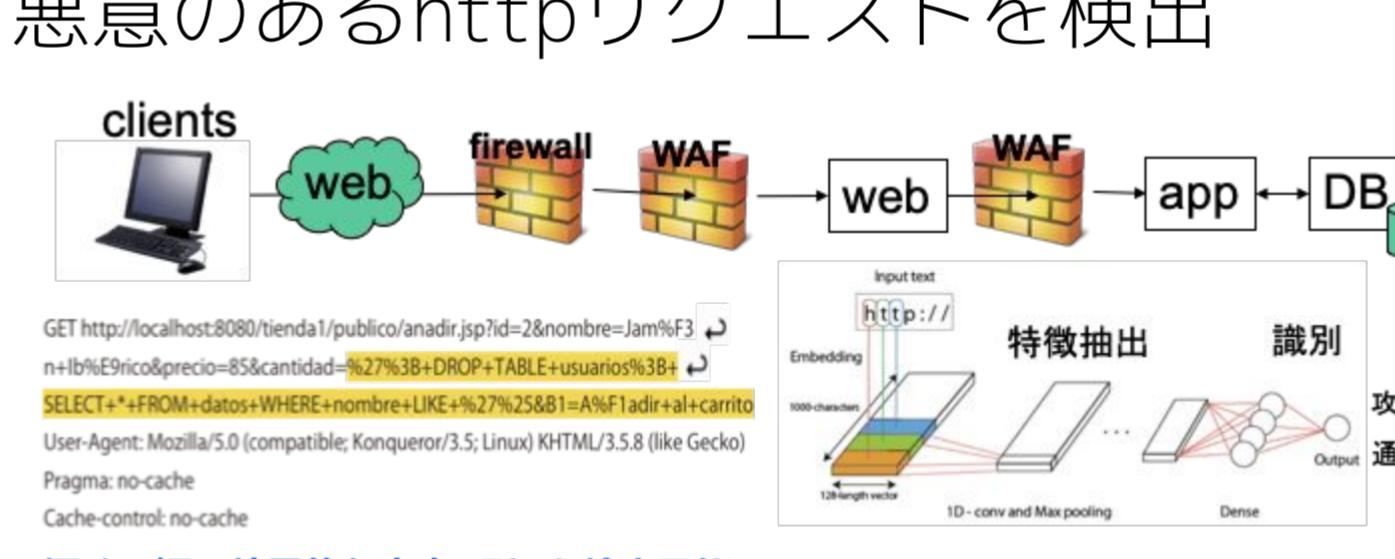
For All
For Women
For Men

■ 深層学習 x セキュリティ

深層学習を用いた Web Application Firewall (WAF) の作成 IEEE CSPA 2018

従来のテンプレートを用いたblacklist方式、垂墨攻撃を防げない

悪意のあるhttpリクエストを検出



HTTP DATASET CSIC 2010 Dataset (normal 36k + attack 25k = 61k)

Detection accuracy: 98.8 % @ 2.35 ms (10 fold-cross valid)

同じデータセットこれまで報告されている最善の結果: 82.0 % with Naive Bayes

深層学習を用いた悪意のあるソフトウェア (malware) の解析 IEEE CSPA 2018

Malware解析: 新種や垂墨の増加が著しく解析が困難に

表層解析: ファイルの特徴を解析

動的解析: 実際に稼働させて動作を解析

静的解析: 機械語や命令を解析

x malware特徴の定義が困難

Character-level CNNを利用

malwareの隠れた特徴を

学習より獲得

- ・ 展開不要 (高速)
- ・ 世界最高レベルの識別精度

Android APKファイル (圧縮ファイル) そのもの

たった1024byteの解析のみで 93 ~ 96% の識別率を達成

■ 悪性黒色腫 自動診断システム

早期発見が大切だが診断が難しいメラノーマ診断支援

(医師の診断精度: 75-84%)

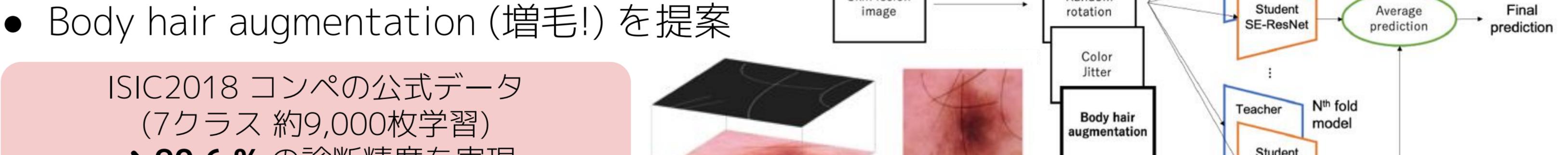
機械学習で最先端の

- ・ SE-ResNet101 (101層CNN) をベースに
- ・ Mean Teachers (半教師あり学習)
- ・ Body hair augmentation (増毛!) を提案

ISIC2018コンペの公式データ

(7クラス 約9,000枚学習)

→ 90.6 % の診断精度を実現



現時点までの対外的成果 (2019年 8月現在)

- ・ 原著論文: 49 件 (うち英文誌33 件)
- ・ 国際会議論文: 69 件
- ・ 国内研究会などでの発表: 100 件+

現在の外部研究資金

- ・ 文部科学省 科学研究費補助金 基盤 C (平成29-31年)
 - 深層学習と次元削減技術を用いた環境変化に
口パストな植物病害自動診断システムの開発
- ・ 農林水産省 受託プロジェクト
 - AIを活用した病害虫診断技術の開発 (平成29-33年)
+ 総務省 add on project PRISM (平成31-33年)

これまで他に

- ・ 科研費6件, JST (科学技術振興機構) の研究助成6件, 企業からの受託・寄附研究 数件

共同研究先

- ・ Johns Hopkins University, The center of Imaging Science, Radiological Science
- ・ University of Central Arkansas., Dept of Computer Science
- ・ Loughborough Univ., Dept of Computer Science
- ・ 東京女子医科大学医学部 皮膚科学教室
- ・ 東海大学医学部 放射線科学教室
- ・ 埼玉県農林総合研究センター (含む 計24府県の農業試験場)
- ・ 国立研究開発法人 農研機構 他企業など

