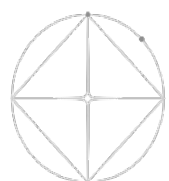


深層学習によるテキストチャ画像処理

電気通信大学 大学院情報理工学研究科
電気通信大学 人工知能先端研究センター
電気通信大学 脳・医工学研究センター

庄野 逸 (shouno@uec.ac.jp)



Shouno Lab.

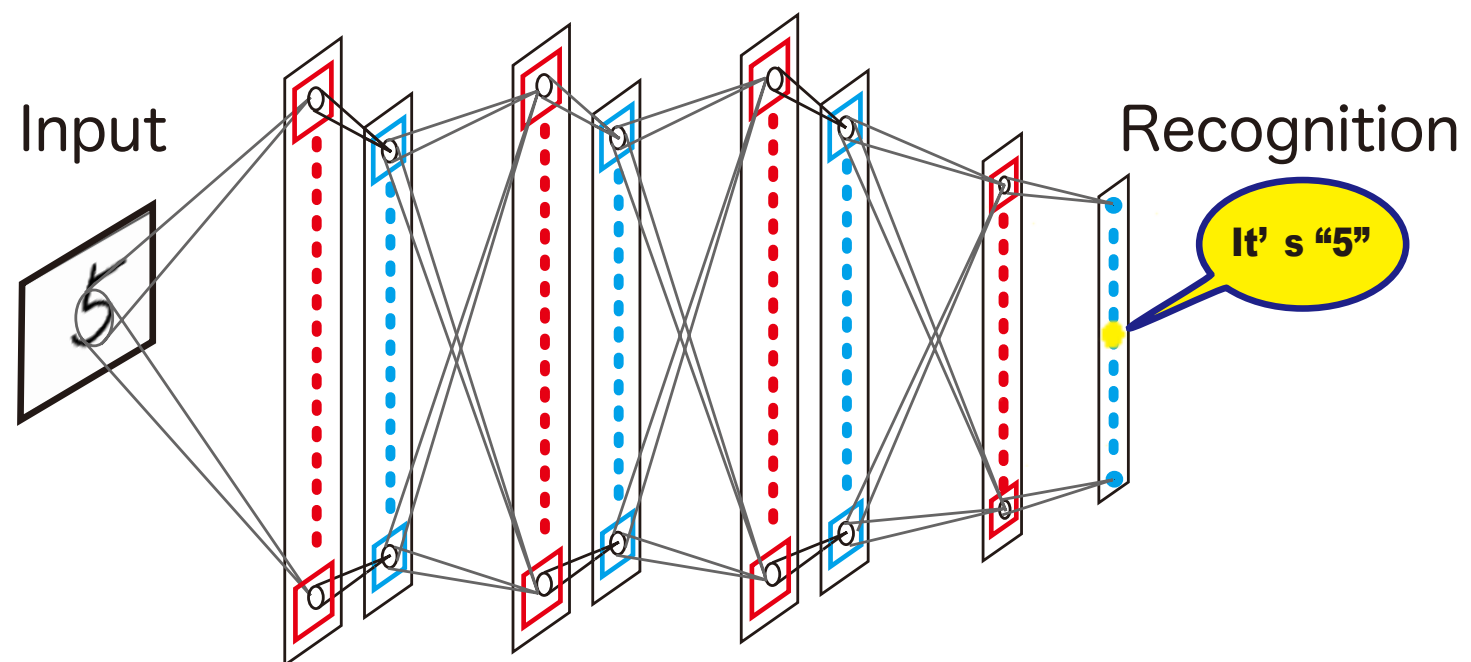
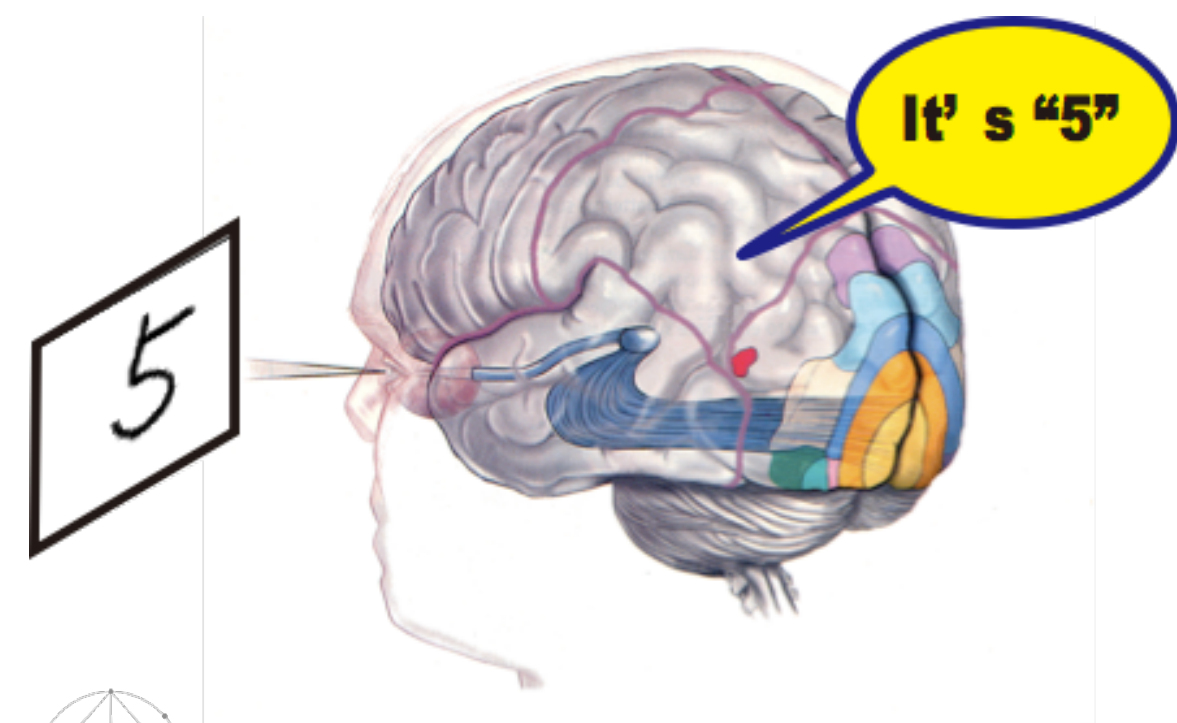


Correspondence and Fusion
of Artificial Intelligence and Brain Science

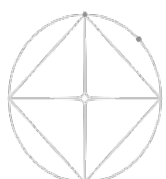


CNN(畳み込みニューラルネット)とは？

- ニューラルネットを用いた機械学習(AI)技術
 - 視覚の機能模した構造で畳み込み演算で記述
 - 深い階層構造を持つことが特徴
- 画像処理，音声処理，テキスト処理の分野で無類の強さを発揮中



Neocognitron/CNN [Fukushima 80]

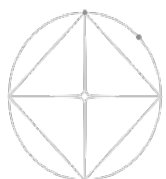


転移学習による CNN への知識埋め込み

Suzuki, Sakanashi, Kido, Shouno, IPSJ TOM 11(3), pp 74-83

<http://id.nii.ac.jp/1001/00192871/>

<https://arxiv.org/abs/1810.06282>

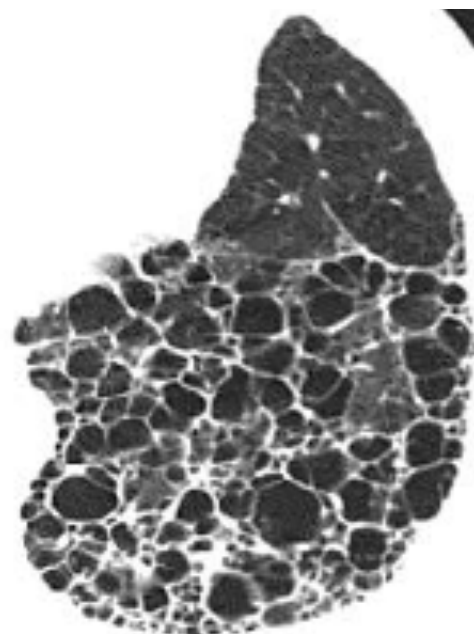


Shouno Lab.



計算機診断支援(CAD)システム

- 計算機診断支援 (CAD)システムの技術基盤形成
 - 計算機支援による医師の意思決定補助
- CAD システムが持つべき特性は何か？
 - 診断性能
 - 適切な診断理由説明
i.e., どの特徴が診断に効くのか？ (特徴選択問題)



direct interpretation



CAD system



second opinion



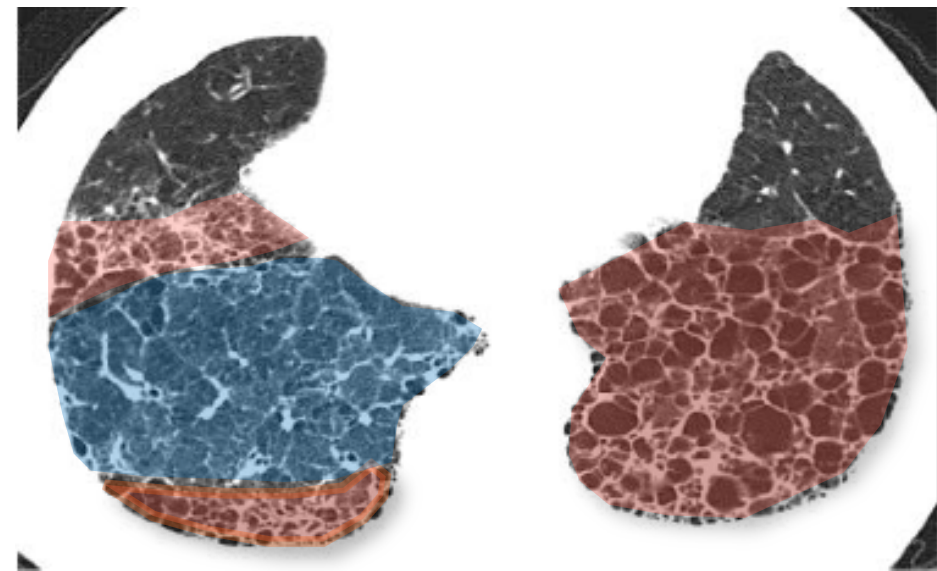
Diagnosis

びまん性肺疾患とその肺CT画像概観

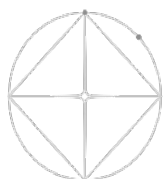
- 肺広範囲に病巣が広がる疾患の総称
 - 肺の線維化を引き起こすことによる呼吸不全
- 早期発見，加療が重要
- Diffuse Lung Disease (DLD) パターンから分類したい
 - 形状ではなく陰影パターン（テクスチャ）の識別



Normal Lungs

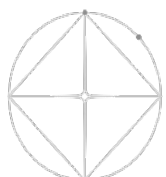
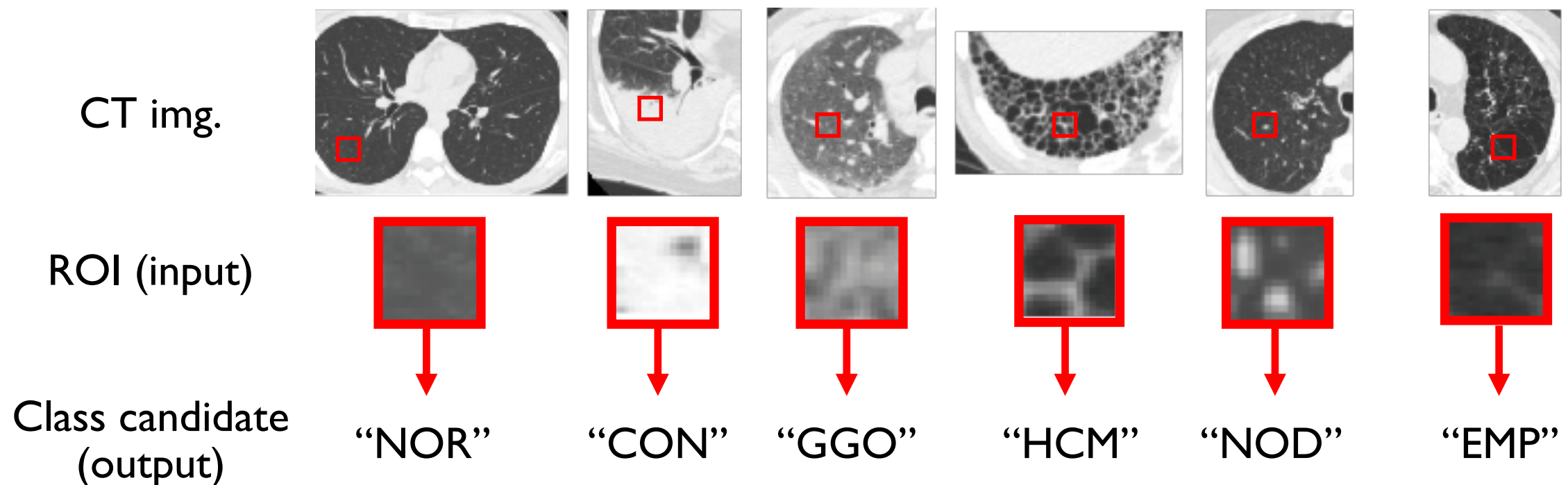


IIP Lungs (HCM, GGO)



パターン認識技術によるCADシステム

- DLD ROI (region of interest) パターン判別機構の確立
- 従来手法: 局所特徴量と識別器の組み合わせ
[Gangeh+10][Sorensen+10][Xu+11]
- 大規模データセットは獲得しにくい
症例数で102 オーダーが限度



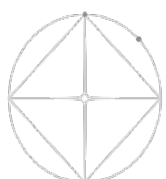
びまん性肺疾患データセット

		Training	Evaluation
Conolication(CON)		143(13patients)	26(14patients)
GGO		609(14patients)	46(14patients)
Reticular(RET)		210(8patients)	66(9patients)
Honeycomb(HCM)		282(10patients)	73(9patients)
Emphysema(EMP)		4406(10patients)	296(11patients)
Nodular(NOD)		762(9patients)	65(10patients)
Normal(NOR)		5371(11patients)	355(11patients)
Total		11783	927

大阪大学病院から提供

ROI サイズ = 32x32 [pixels] ~ 16 x 16 [mm²]

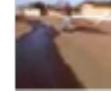
頑張ってROIサンプルを取り出しても 1.1×10^4 個程度



Shouno Lab.



深層学習のパラメータ数とデータ数

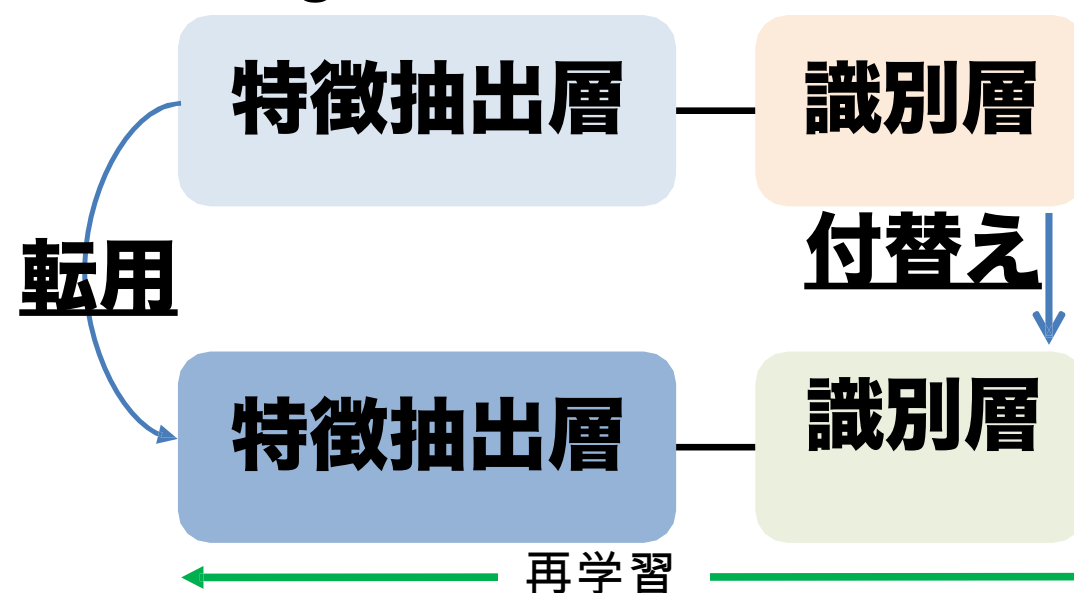
	LeCun+95	Fukushima+98 Fukushima+13	Krizhevsky+12	Simonyan+15	Ours [Krizhevsky+12]
Data type	Handwritten digits (MNIST) 	Handwritten digits (ETL-1) 	Natural images (ImageNet) 	Natural images (ImageNet) 	DLD ROI 
DB size	6×10^4	1.0×10^4	1.2×10^6	1.2×10^6	1.1×10^4
#layers	6	13	13	19	13
#weights	1.0×10^5	6.9×10^5	1.3×10^7	1.4×10^7	1.3×10^7

- NN を訓練するには**十分な数の学習データ**が必要
 - モデルの複雑度は比較的高い
自由パラメータ 10^7 オーダー
 - 深層モデルを十分に学習させるには自由パラメータの1/10 の
学習データが必要[Bengio+14]
- 学習データが足りなそうなので **「知識」** を埋め込む

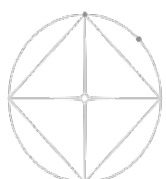
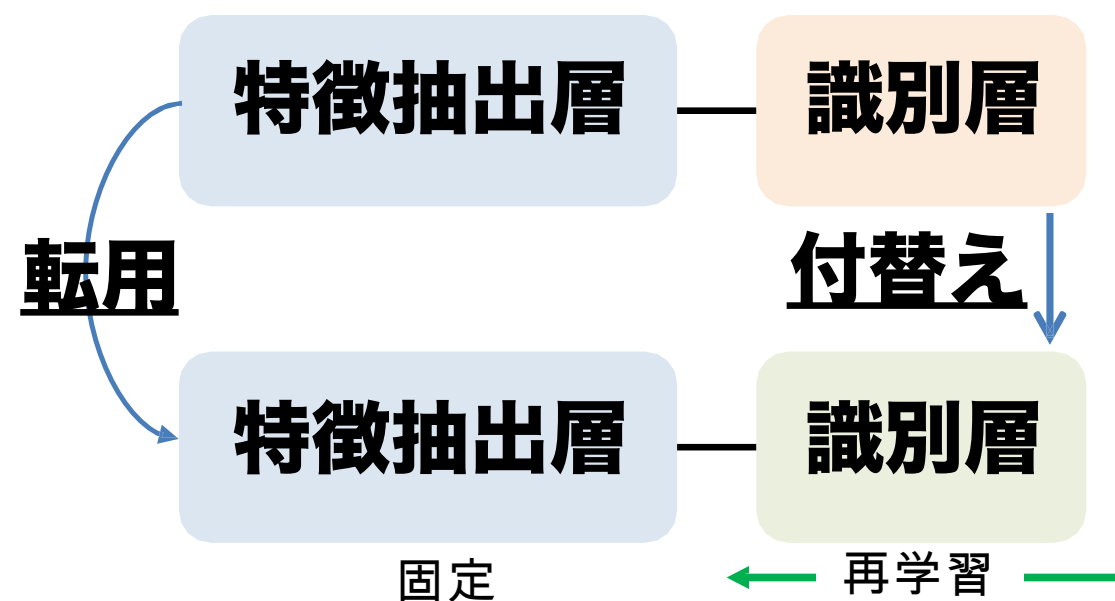
転移学習

- 転移学習 (Transfer Learning)
 - 目的とは違う問題で得た知識を転用するテクニック
- CNNにおける転移学習(Fine tuning)
 - 特徴抽出の性能向上のため特徴抽出層を転用
識別層は新たに初期化しネットワーク全体を再学習

Fine Tuning



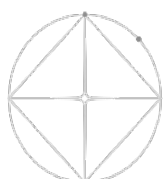
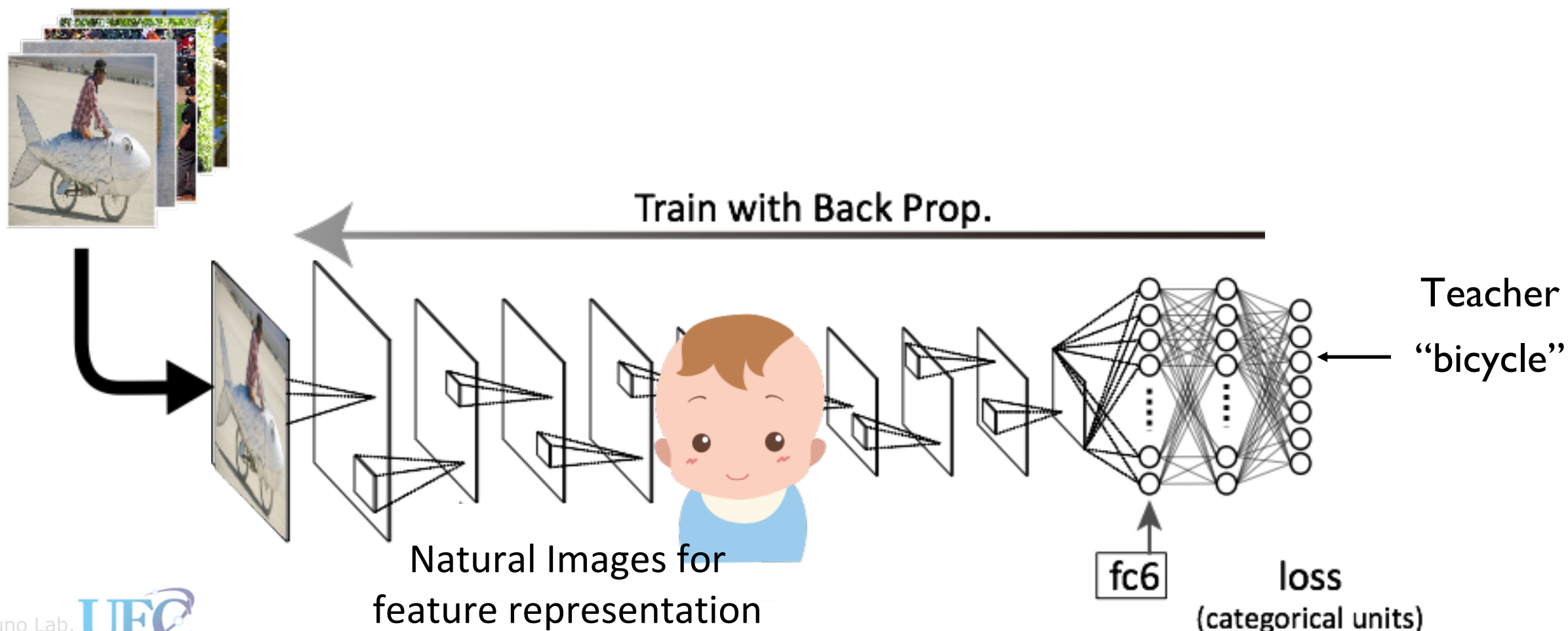
Feature-Extraction



小規模 DB で CNN を学習させるには？

□ 仮定したこと:[Shouno+15]

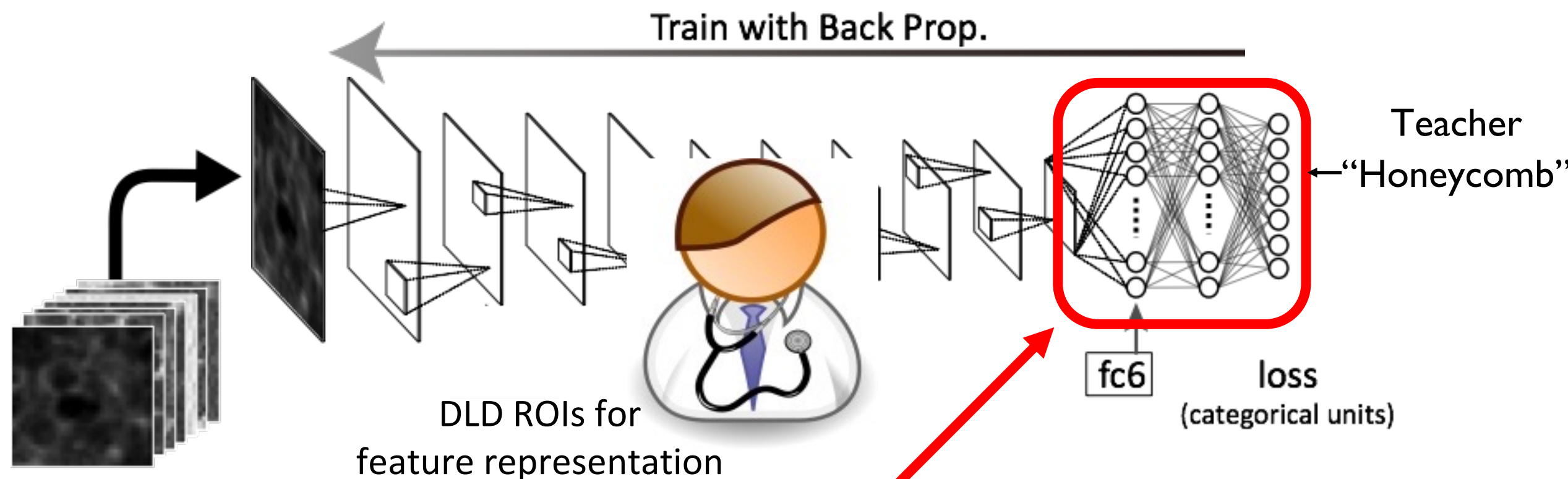
- 画像は何らかの共通構造を持っていて, CNN はその構造を表現できそう
- 自然画像はDLDパターンの構造表現に役立つかも
(特殊な画像は大人になってから学習)



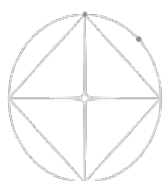
小規模 DB で CNN を学習させるには？

□ 仮定したこと:[Shouno+15]

- 画像は何らかの共通構造を持っていて, CNN はその構造を表現できそう
- 自然画像はDLDパターンの構造表現に役立つかも
(特殊な画像は大人になってから学習)

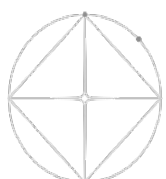
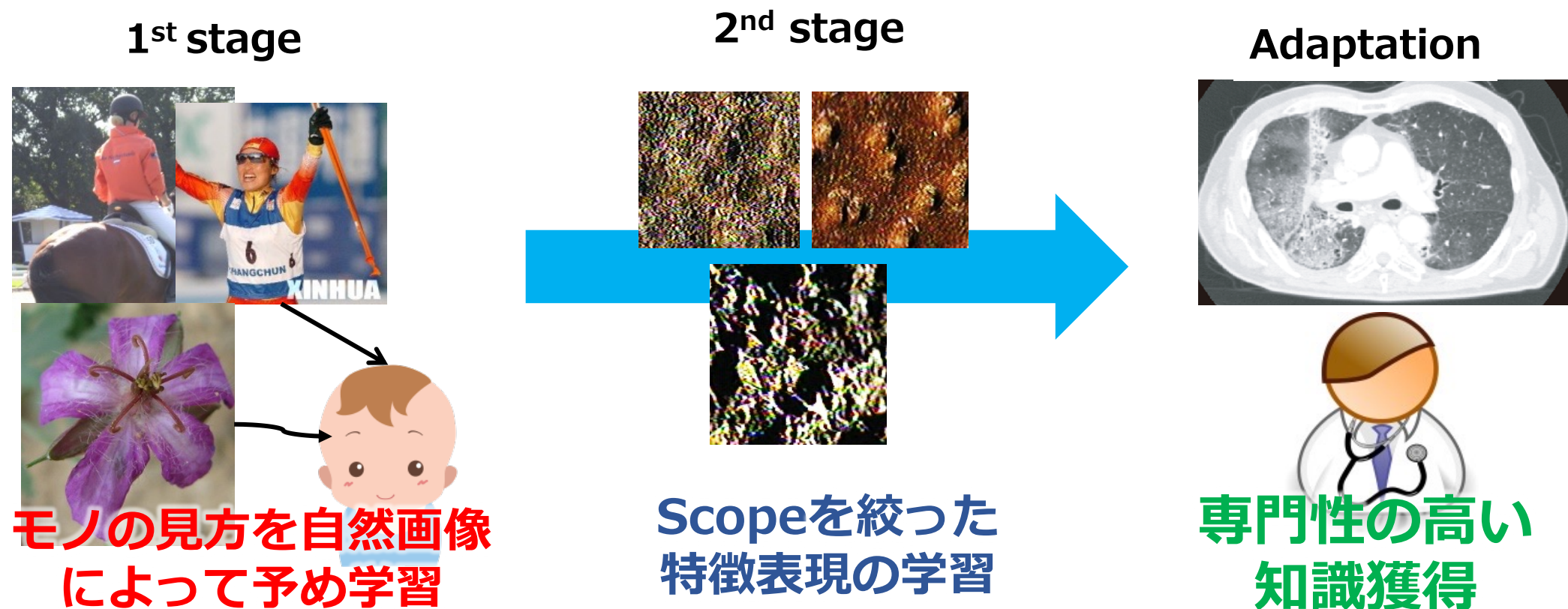


識別器部分を交換



転移元は自然画像（だけ）でよいのか？

- 自然画像を用いた転移学習 [Shouno+15], [Gao+16]
 - 自然画像を Source とした転移学習による性能向上
 - Source と Target が大きく異なっているので本当に良いかは不明 [Azizpour+16]
- 2段階転移学習



Source Domain の選択

□ 自然画像: ISLVR2012 dataset

□ 大規模画像データセット ImageNet のサブセット

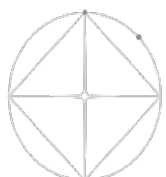
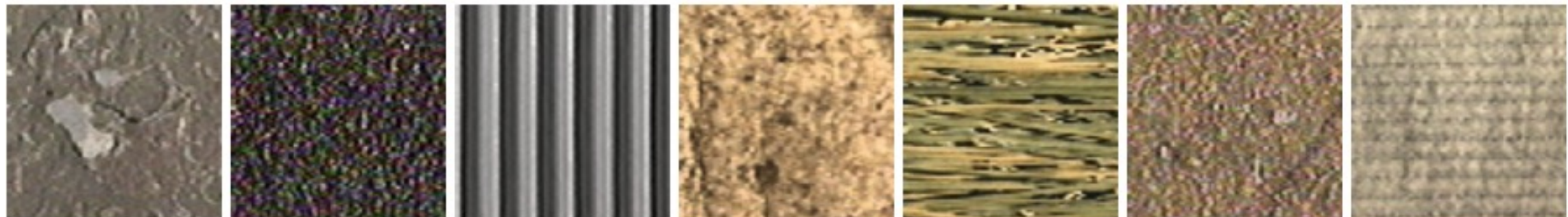
□ 一般物体 1,000 クラスの写真: 1.2×10^6



□ テクスチャ画像: CUReT database

□ Utrecht Univ. が提供するテクスチャ画像セット

□ 61クラスの方法表面写真 1.4×10^4 枚



比較対象

□ 学習過程の異なる4つのモデルの比較

(1) CT画像のみで学習

(2) テクスチャ画像 → CT画像

(3) 自然画像 → CT画像

(4) 自然画像 → テクスチャ画像 → CT画像

← 転移学習なし

← 1段階転移学習

← 2段階転移学習

自然画像

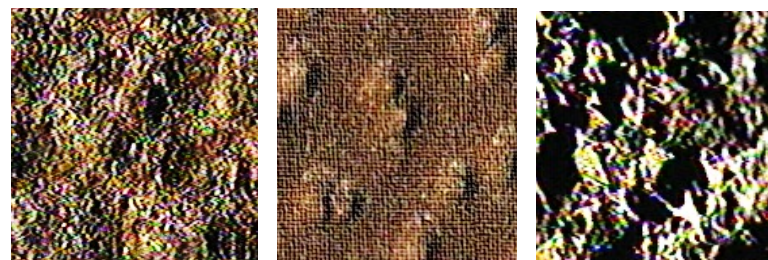
ILSVRC2012



計 1,200,000枚

テクスチャ画像

CURet texture DB



計 14,406枚

CT画像

肺断面の胸部X線CT

(※識別対象)



計 12,888枚

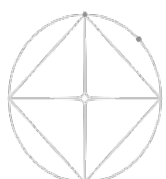
識別性能評価

- 2段階転移学習(4)は全ての性能指標で最も高い性能
- テクスチャ画像からの転移(2)は, (1) よりも性能悪化

転移学習	なし (1)	テクスチャ (2)	自然画像 (3)	2段階 (4)
Accuracy	0.9277	0.9201	0.9558	0.9601
Recall	0.9583	0.9412	0.9484	0.9739
Precision	0.9590	0.9417	0.9471	0.9719
F1-Score	0.9583	0.9411	0.9470	0.9724



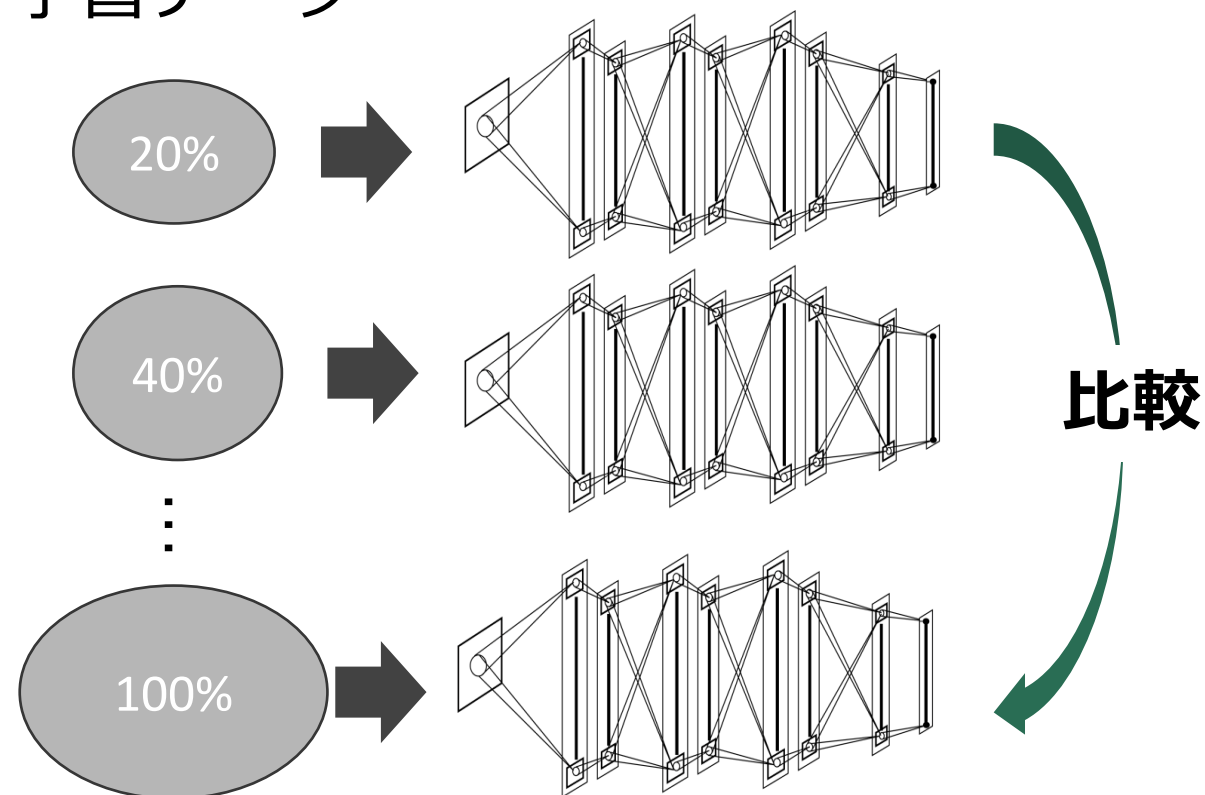
Source domain選択の重要性を示唆



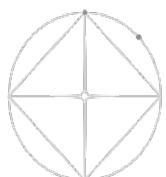
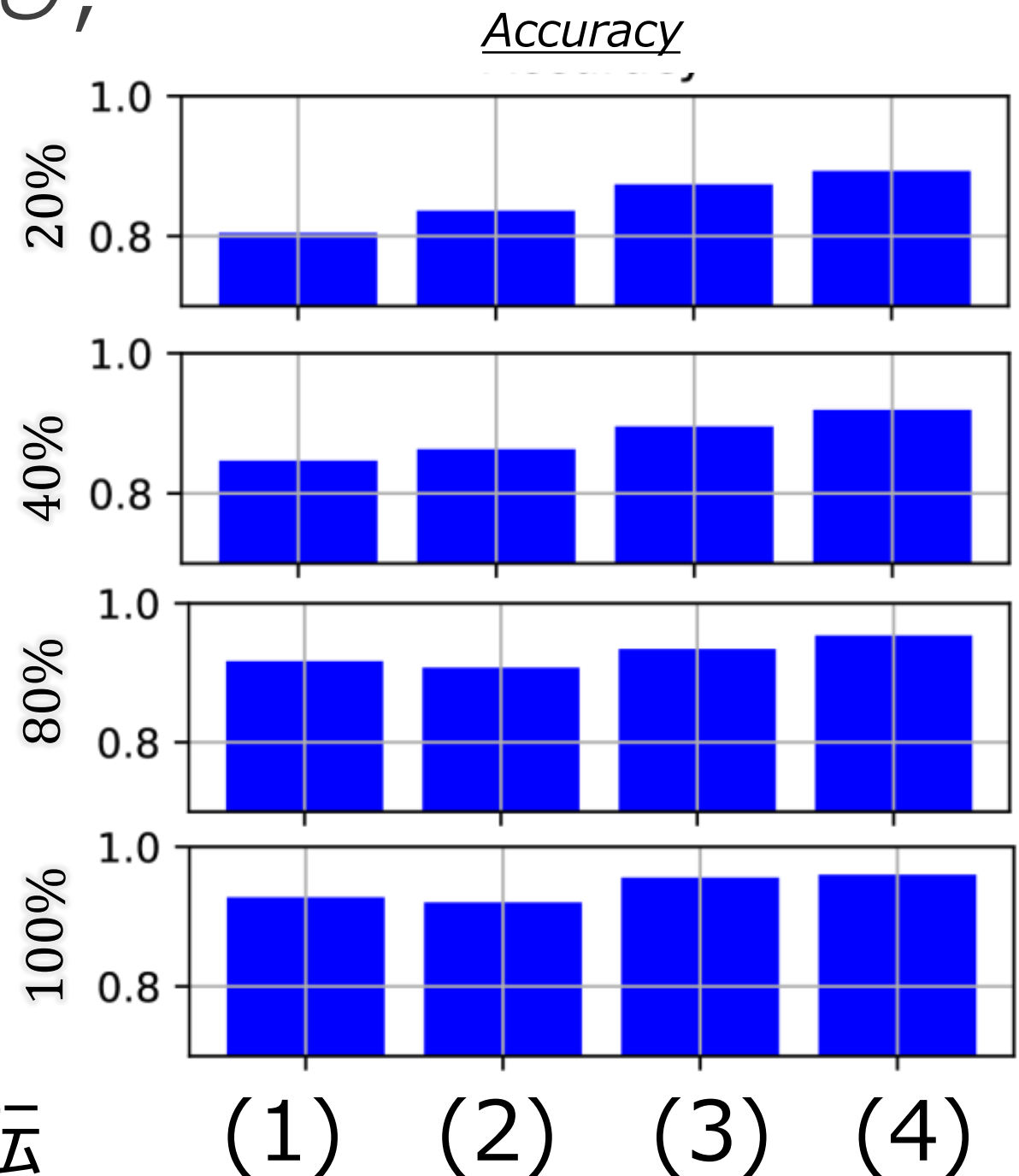
頑健性評価

- 学習データ数をコントロールし、性能劣化の度合いを調べる

学習データ



- ✓ (4) は最も頑健
- ✓ 減少比が大きいと(1),(2)が逆転



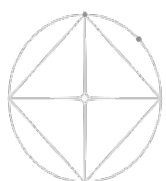
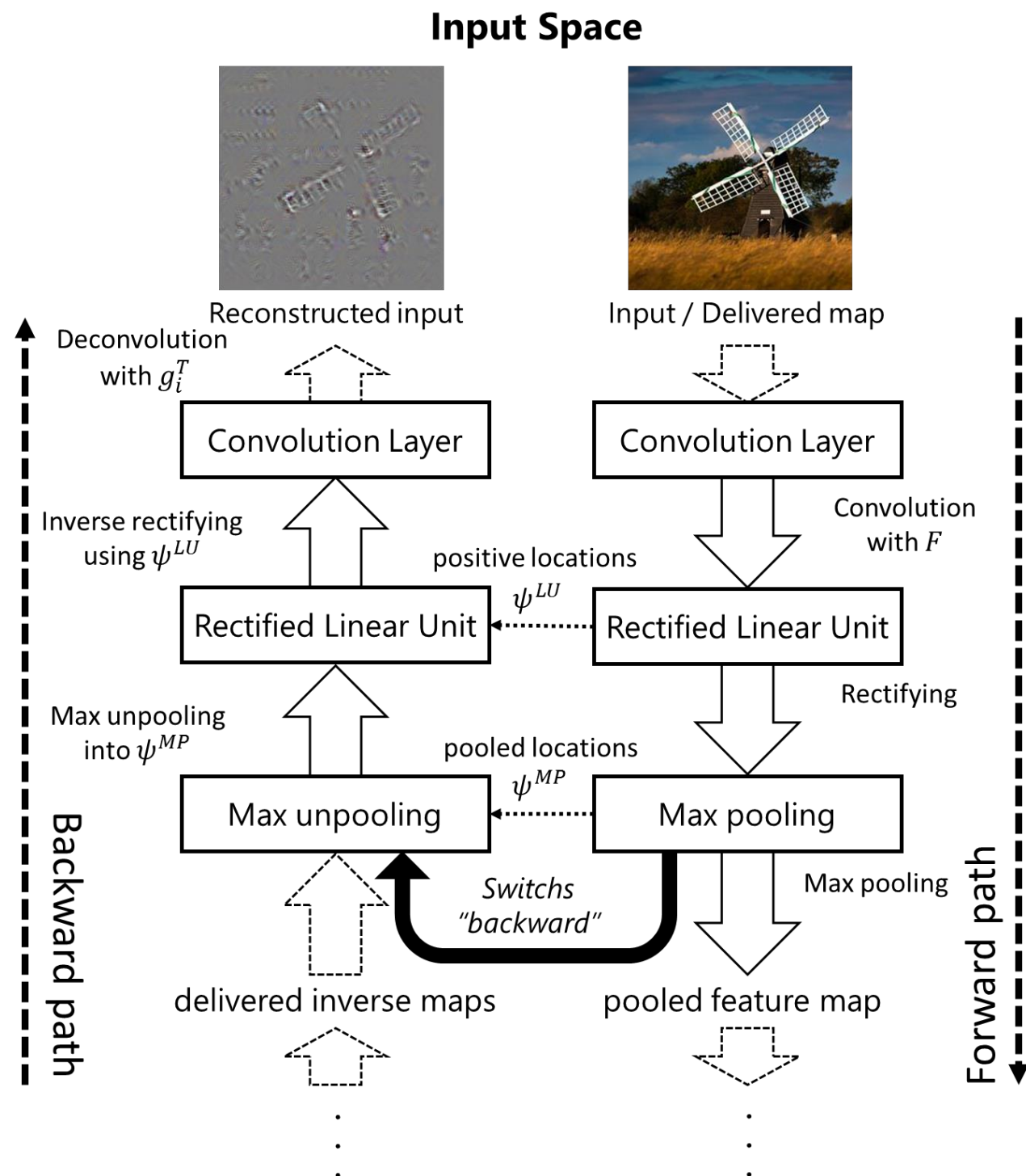
DeSaliNet による可視化

- [Mahendran+16]による特徴表現の可視化手法
- 入力の**どういった成分**が特徴表現に寄与するかを入力空間の画像として観察
- DCNNを一つの合成関数と捉え、要素ごとの逆変換で特徴表現を入力空間に復元

$$f = g_1 \circ g_2 \circ \dots \circ g_n$$

$$\Updownarrow$$

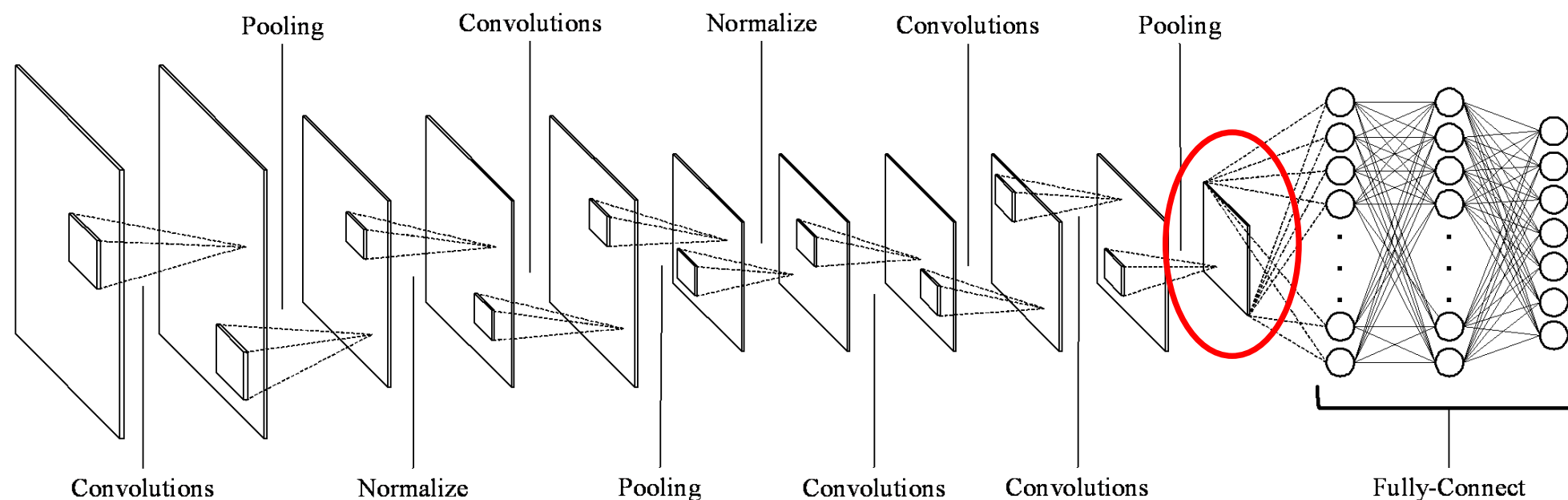
$$f^\dagger = g_n^\dagger \circ \dots \circ g_2^\dagger \circ g_1^\dagger$$



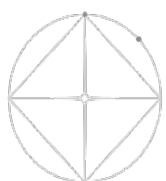
特徴の可視化

なぜ転移学習によって識別性能が向上するのか？

➡ 識別層に入力される特徴表現の変化に着目

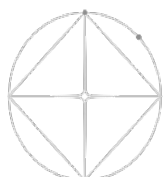
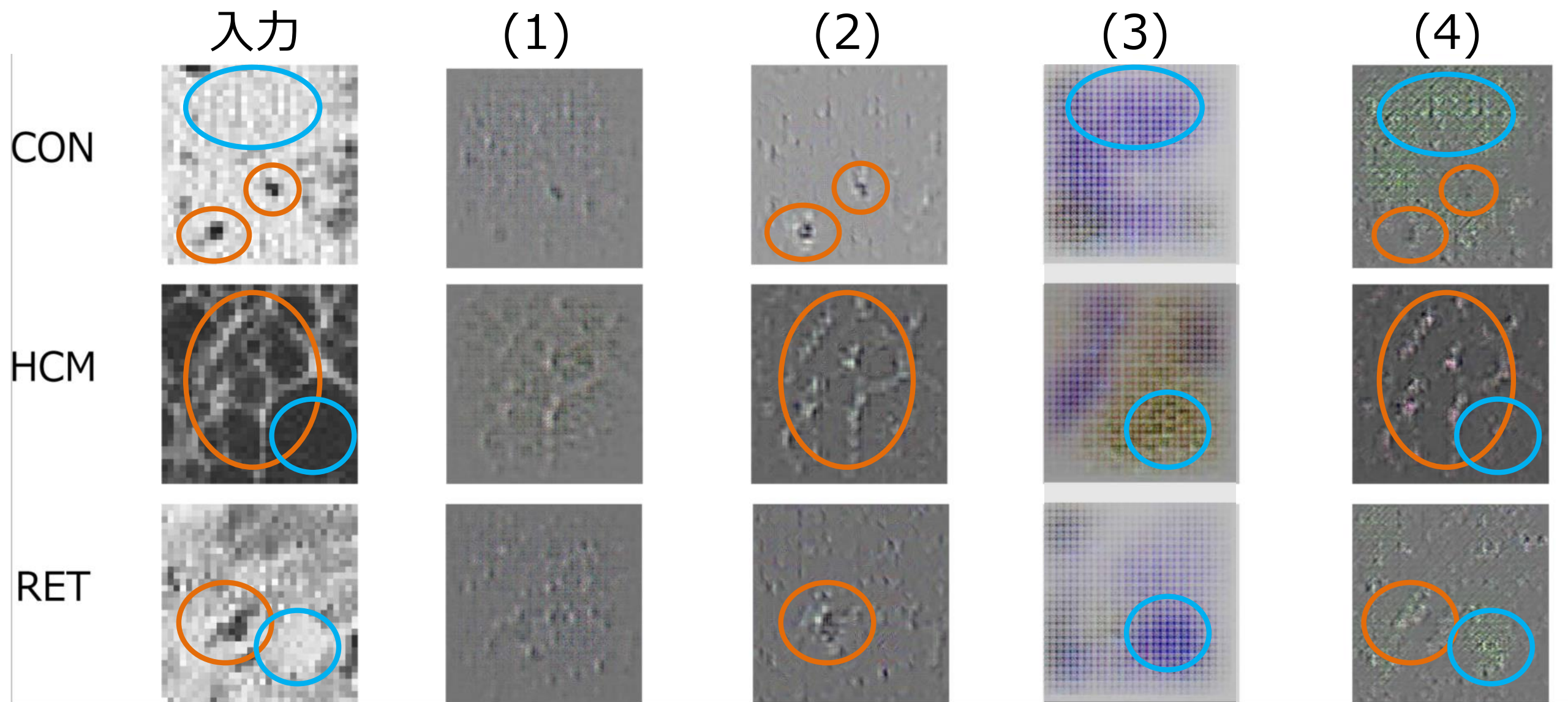


転移元ドメインによる特徴表現の変化を観察することで
DCNNにおける転移学習のメカニズムを理解する



特徴表現の可視化結果

- 自然画像を学習した(2) はエッジ構造に，テクスチャのみの(3) は模様に， 2段階のもの(4) は双方に活性



まとめ: 深層学習によるびまん性疾患診断

- CAD技術に必要な識別性能の向上を試みた
- 手法としては深層学習 + 転移学習を用いて構築
 - データが少ない分野において「知識」を埋め込む方法が必要
 - 可視化により特徴の位置に関しては推測可能
- 良好な識別性能, 頑健性を得られることを確認
 - 精度評価で 96% 程度, 頑健性においても最も良好
- ただし, 転移元の知識をどのように設計するかは, 検討が必要
 - テクスチャデータで事前学習した場合は, 精度低下を引き起こす.

Future work

- 画像から画像を作る深層学習は構成可能だが、中間層表現に「知識」を与えないと多分つらい
- RTE解—観測モデルのマッピング取得を Source とする
- 転移学習による性能向上ともっともらしいテクスチャ埋め込みを考えることが多分必要

