

深層学習を用いた Digitally Reconstructed Radiographsに 基づく変形性股関節症の多指標自動分類

○ 梶田 修慎^{§1} 崇風 まあぜん^{§1} 大竹 義人^{§1} 上村 圭亮^{§2}

高尾 正樹^{§3} 菅野 伸彦^{§2} 佐藤 嘉伸^{§1}

^{§1}Nara Institute of Science and Technology, Division of Information Science

^{§2}Osaka University, Graduate School of Medicine

^{§3}Ehime University Graduate School of Medicine

背景: Hip OA・グレーディング指標

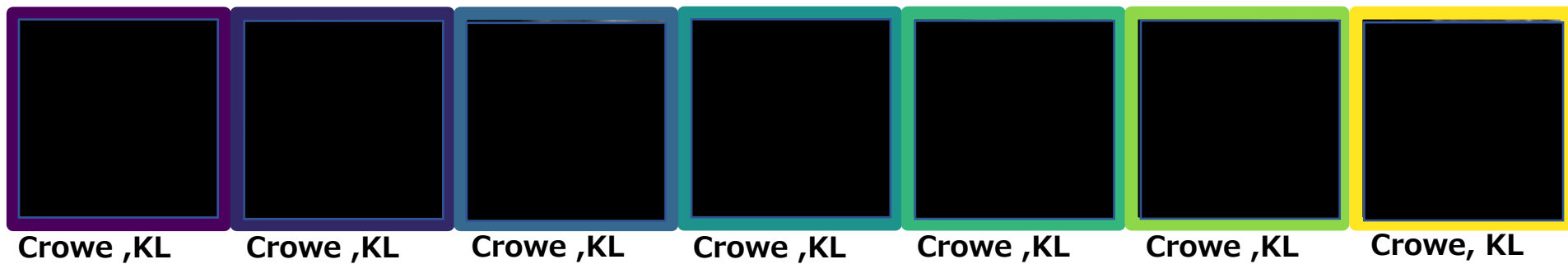
2

- 変形性股関節症(Hip OA)は現代の超高齢化社会において問題視されている疾患
 - 疾患進行の評価には整形外科医の臨床専門性が必要
- 評価にはKL(Kellgren&Lawrence)[*](Crowe分類[**])が用いられる
 - 疾患の重症度に応じたグレードが付与される
 - Hip OAの患者は筋力低下を経験する[***]

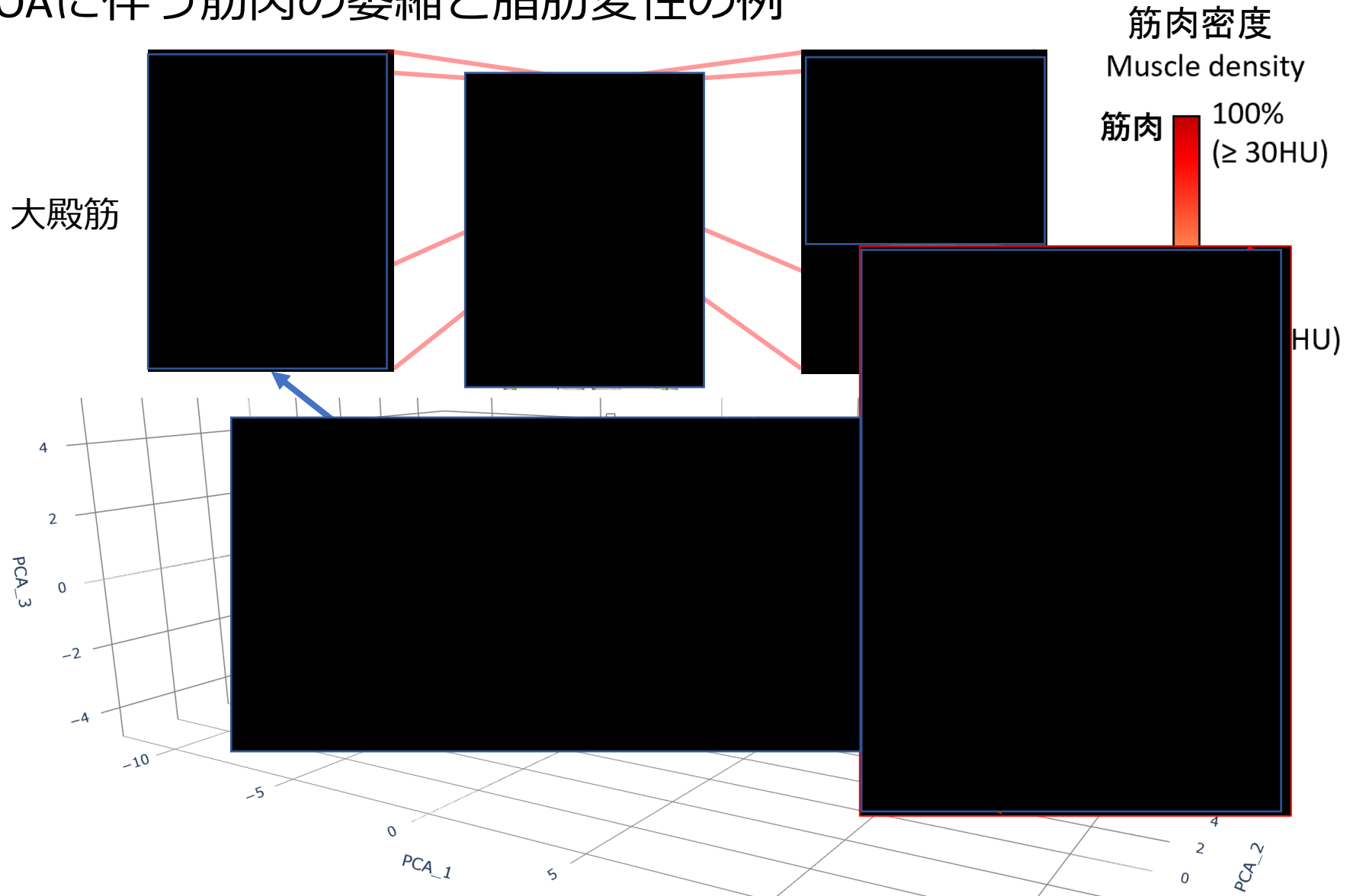
正常

Hip OAの進行とグレーディング

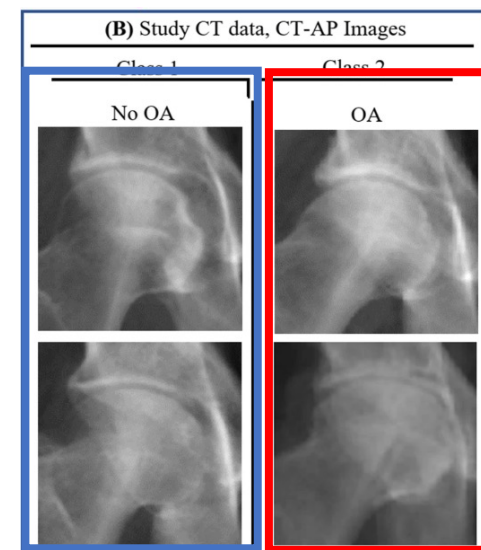
重症



- OAに伴う筋肉の萎縮と脂肪変性の例



- 近年、畳み込みニューラルネットワーク(CNN)を用いた自動分類手法が報告された
- Gebre et al. Osteoporosis International, 2021
X線やCT画像から作成した疑似X線画像(DRR)を用いた
結果 (DRR) : 94症例, Accuracy : 83.3%
- Xue Y, et al. PlosOne, 2017
事前学習済みのVGGを用いて、両側の股関節を含めた
X線画像に対してHOA
結果 : 420症例, Accuracy : 92.0%



二値分類の問題点：臨床的に重要な形状や輝度値の変化を評価することができません

- HOAを診断する指標としてOAの重症度を示すKLに加え、脱臼度を示す指標であるCrowe分類を組み合わせ、疾患の進行具合を正確に把握可能な自動分類器を開発する
- 不確実性を評価することで、アノテーションが施されていない数千症例の大規模なデータベースに対してのアクティブラーニングが可能か示す

手法

■ モデル

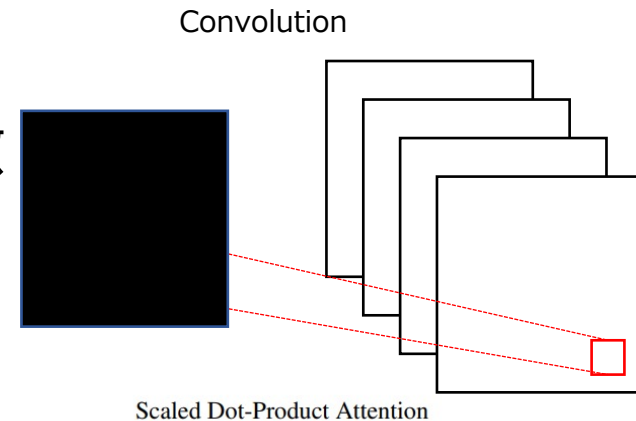
VisionTransformer_Base16 (Attention機構を用いたモデル) [*]

VGG16 (畳み込みを用いたモデル) [**]

DenseNet161 (畳み込みを用いたモデル) [***]

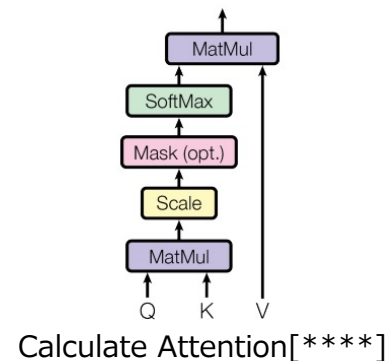
■ 畳み込みニューラルネットワークの特徴

- Kernelを移動して畳み込み演算を行う
- 局所的な情報の参照に優れている



■ Attentionを用いたネットワークの特徴

- 畳み込みを用いずに複数回にわたって Attention(注目度)を計算する
- 大域的な情報の参照に優れている



[*] A. Dosovitskiy, L. Beyer, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, ICLR 2021

[**] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, ICLR 2015

[***] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

[****] A. Vaswani, N. Shazeer, et al. Attention Is All You Need, NIPS 2017

実験

データセット

9

- DRRデータセット

| | |
|--------------------------|---------------------------|
| Image Size | 150 × 150 [pixels] |
| Number of Classes | 7 |
| Number of Images (Cases) | 396 (198 cases) |
| Institution | Osaka University Hospital |

正常

重症

Class 1

Class 2

Class 3

Class 4

Class 5

Class 6

Class 7

Crowe ,KL

Crowe ,KL

Crowe ,KL

Crowe ,KL

Crowe ,KL

Crowe ,KL

Crowe ,KL

| Parameter | Value |
|------------------|---------------------------------------|
| Environment | Pytorch + Torchvision |
| Initialization | ImageNet database pre-training |
| Epochs | 200 |
| Cross-validation | Stratified group 4-fold (15 patterns) |
| Image Size | 224×224 [<i>pixels</i>] |

■ データ拡張

- Resize : 画像サイズを変更する
- Rotate : 画像を回転する
- Blur : ぼかしを加える
- Random Brightness/Contrast : ランダムでコントラストを変化する
- Coarse Dropout : 画像に不規則なマスクをかける

結果

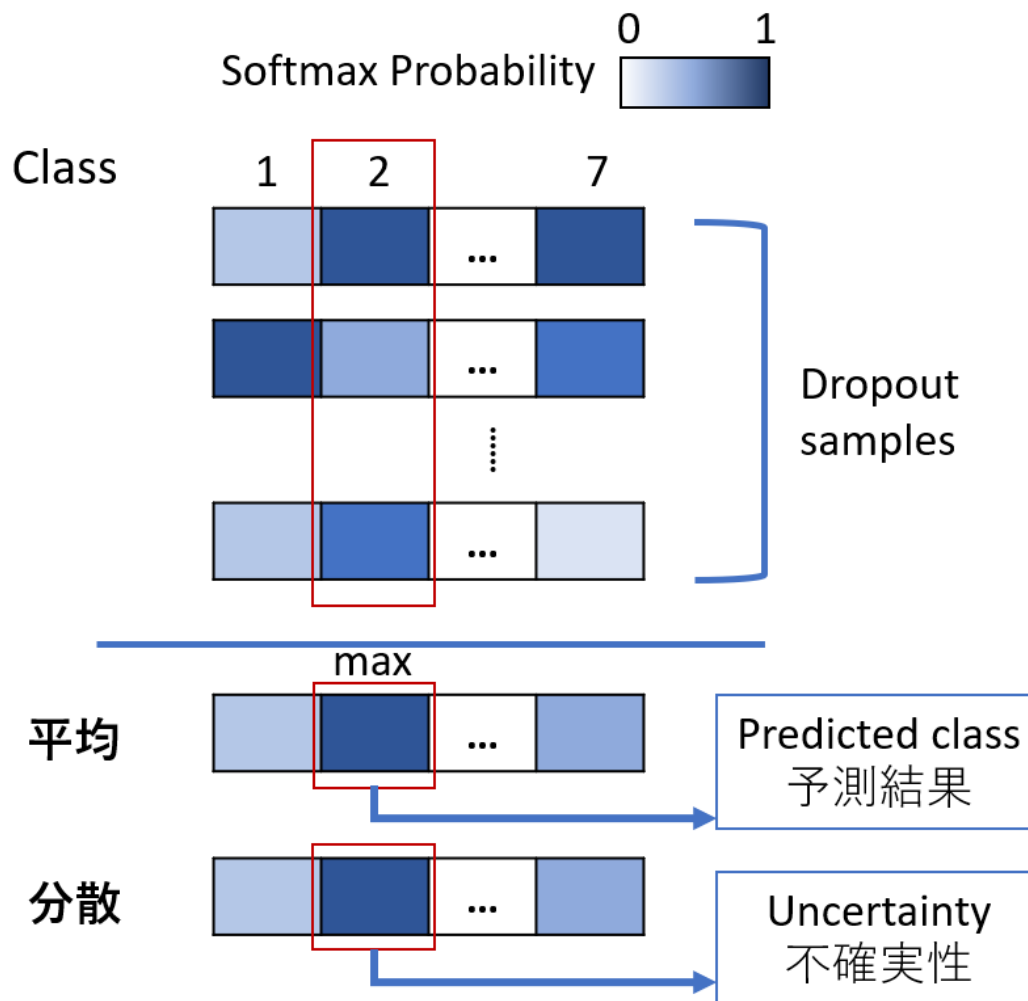
- **テスト時Dropoutありの50回の推論で得られる**
Softmax Probabilityの分散を不確実性の指標とする[*]

$$\frac{1}{N} \sum_{i=1}^N (P_i - \hat{P})^2$$

N : Number of Samples

P : Softmax Probability

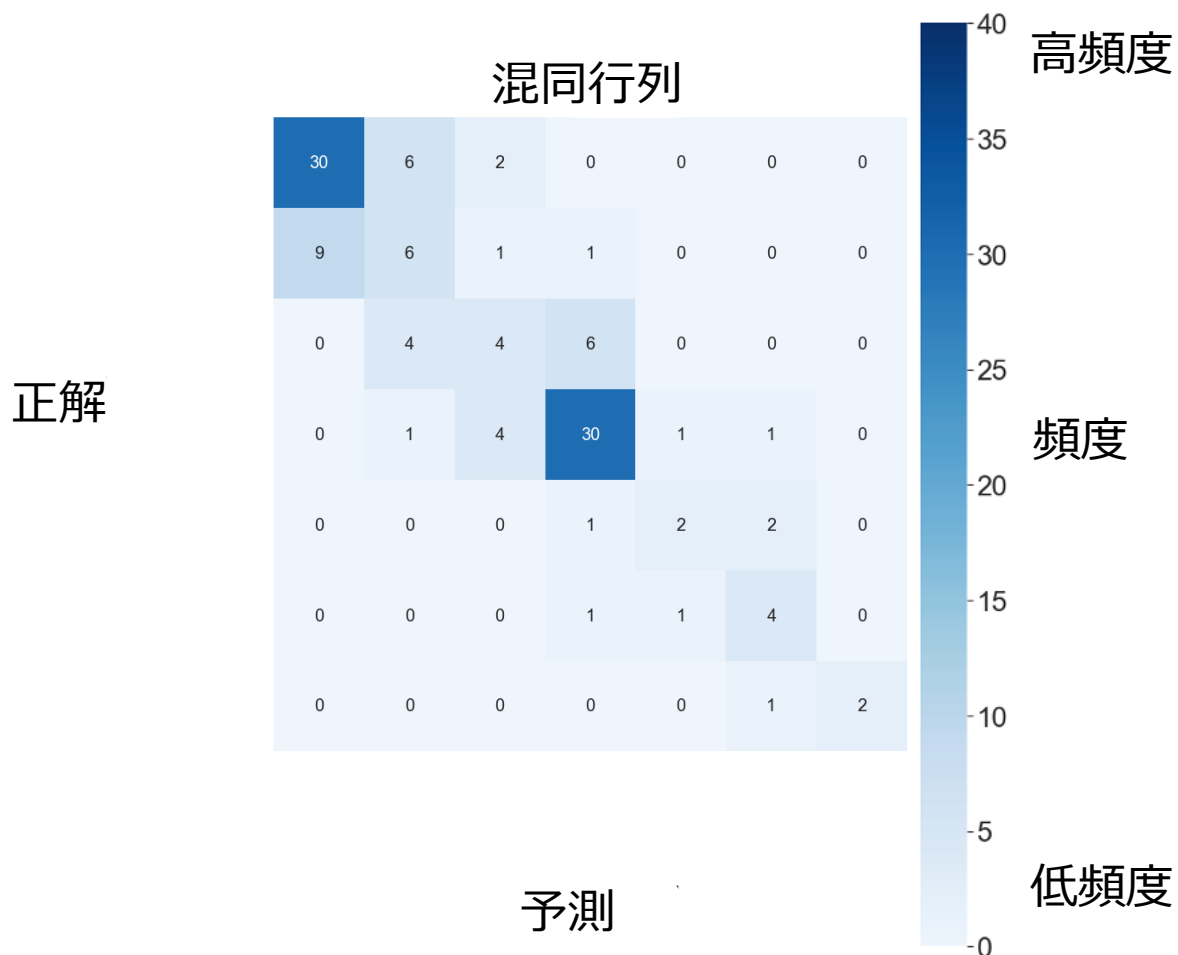
\hat{P} : Average Softmax Probability



[*] Y. Gal, Z. Ghahramani, Dropout as a Bayesian approximation: Representing model uncertainty in deep learning, ICML-16, 2015

■ Accuracy

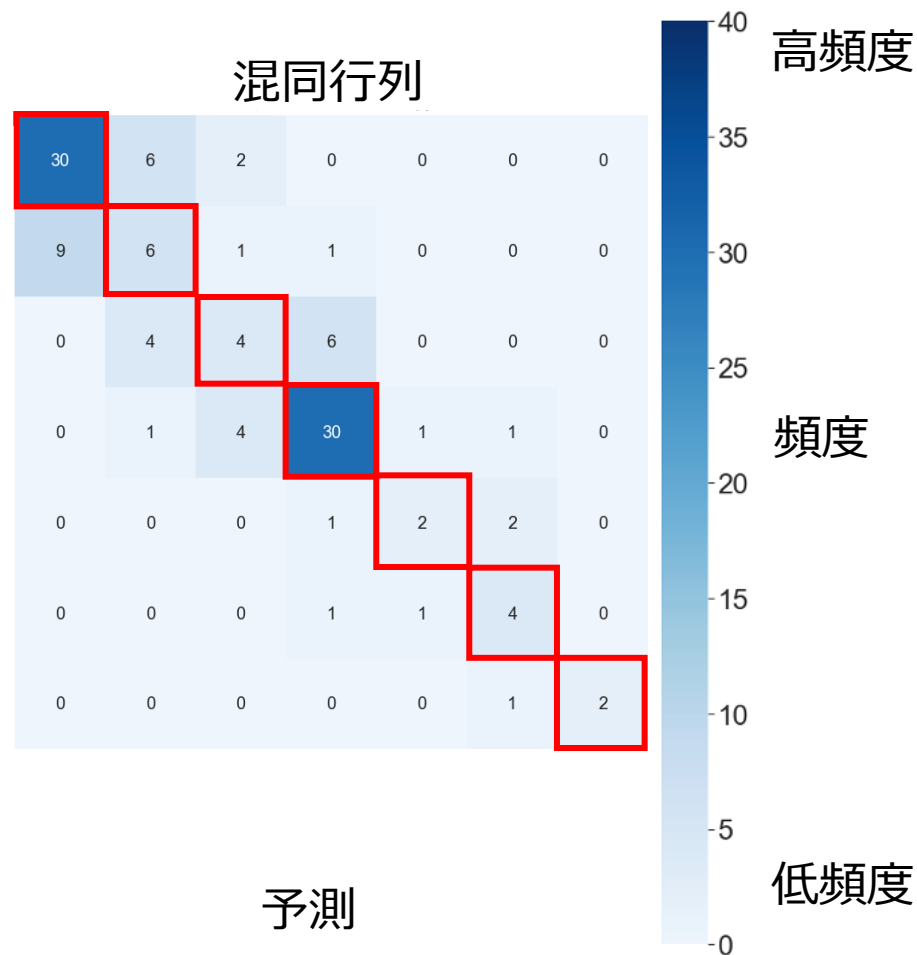
- 予測クラスと正解クラスが一致
- 1クラス分間違っても許容
- 上記以外



■ Accuracy

- 予測クラスと正解クラスが一致 → Exact Class Accuracy
- 1クラス分間違っても許容
- 上記以外

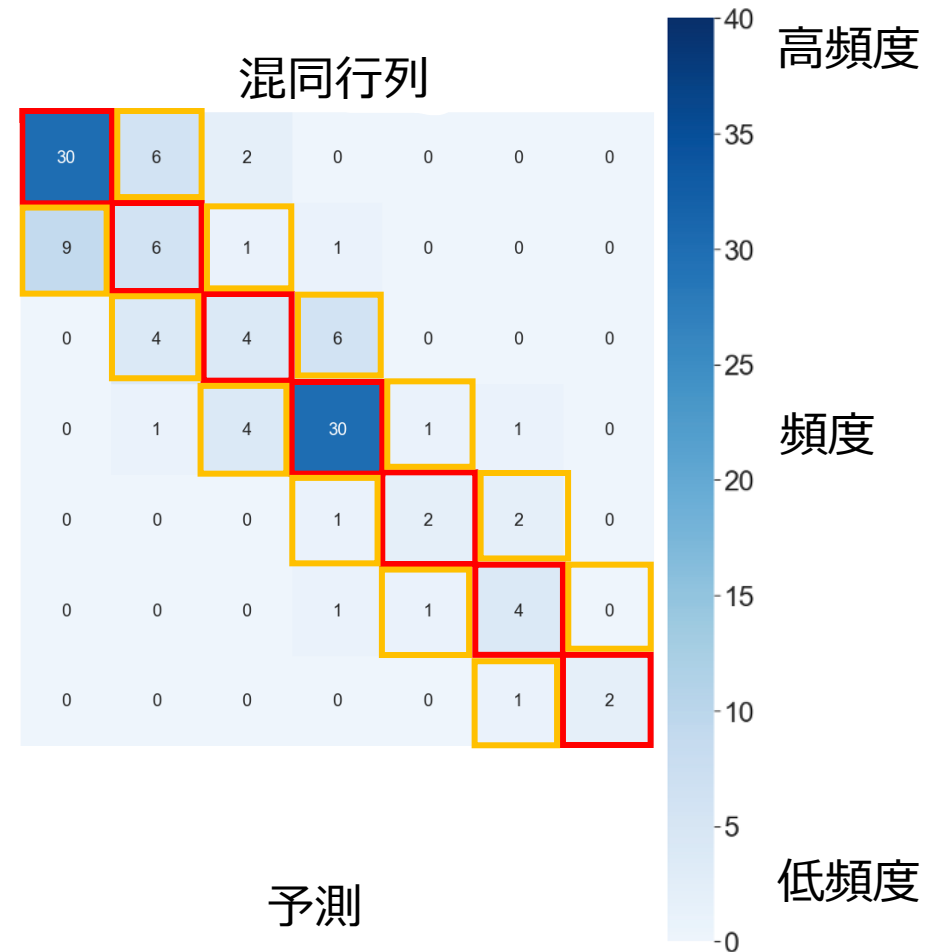
□ Exact Class Accuracy 正解



■ Accuracy

- 予測クラスと正解クラスが一致 → Exact Class Accuracy
- 1クラス分間違っても許容 → 1-Neighbor Class Accuracy
- 上記以外

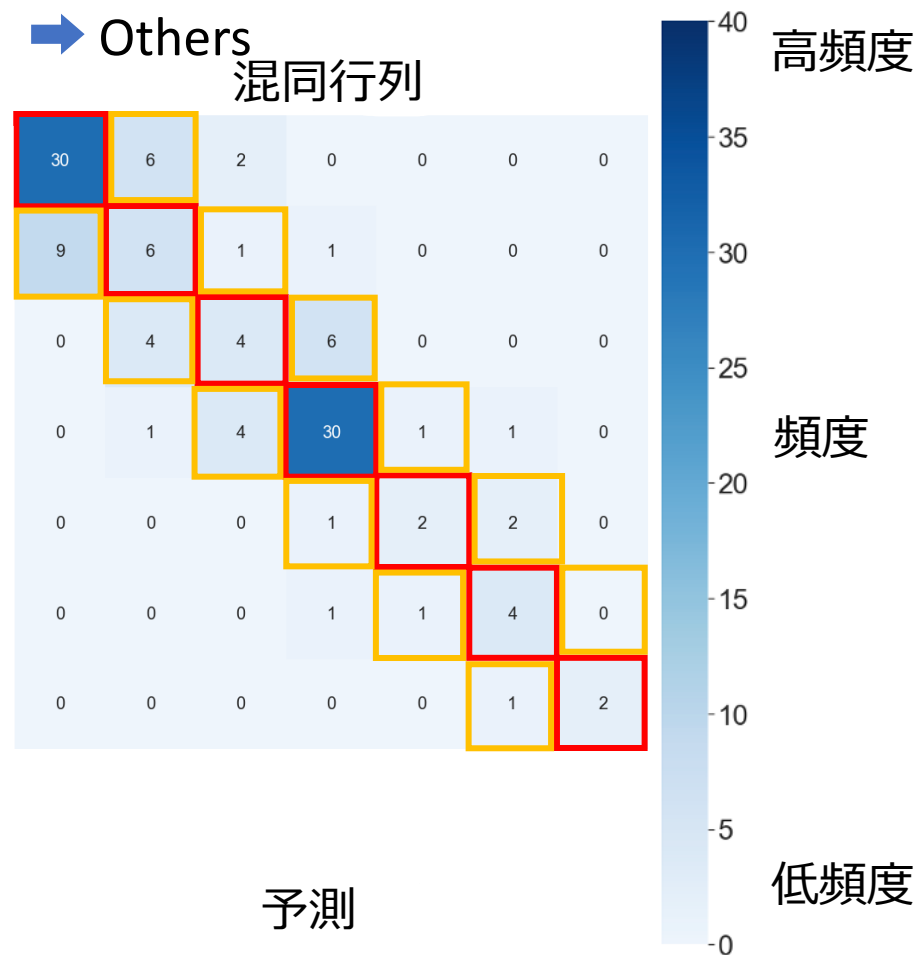
□ Exact Class Accuracy 正解
□+□ 1-Neighbor Class Accuracy



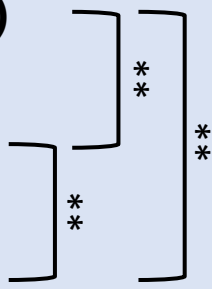
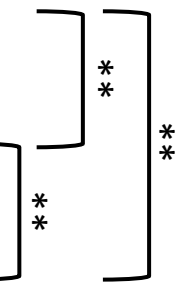
■ Accuracy

- 予測クラスと正解クラスが一致 → Exact Class Accuracy
- 1クラス分間違っても許容 → 1-Neighbor Class Accuracy
- 上記以外 → Others

Exact Class Accuracy 正解
 1-Neighbor Class Accuracy



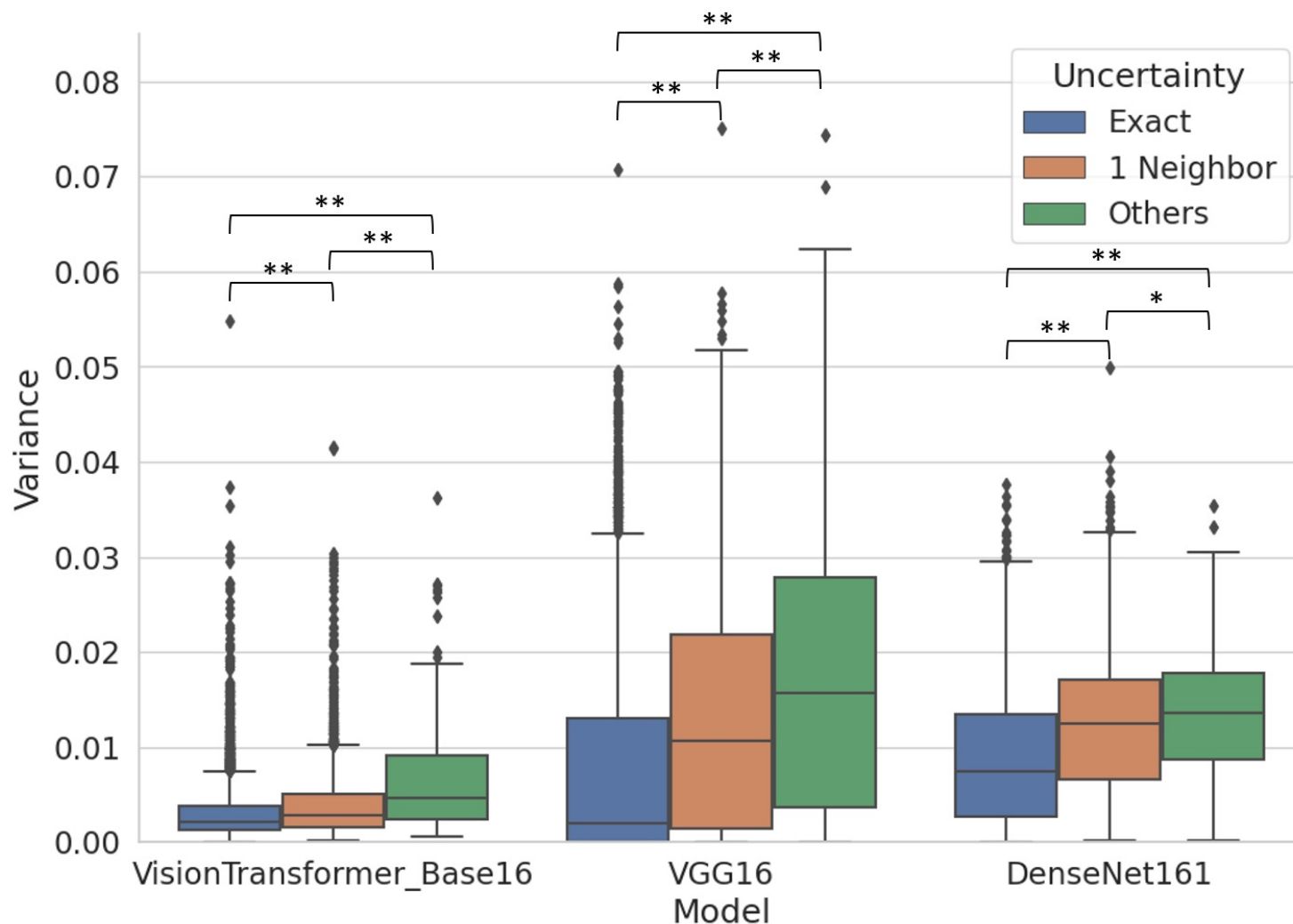
- VisionTransformerが最も良い精度を得られた

| 推論精度(±標準偏差) | |
|-------------|--|
| | Exact Class Accuracy |
| ViT | 0.648(±0.129) |
| VGG | 0.637(±0.148) |
| DenseNet | 0.622(±0.154) |
| |  |
| | 1-Neighbor Class Accuracy |
| ViT | 0.953(±0.131) |
| VGG | 0.947(±0.161) |
| DenseNet | 0.944(±0.160) |
| |  |

** : $p < .01$ (Mann-Whitney U test)

不確実性の評価結果 (Softmax Probabilityの分散)

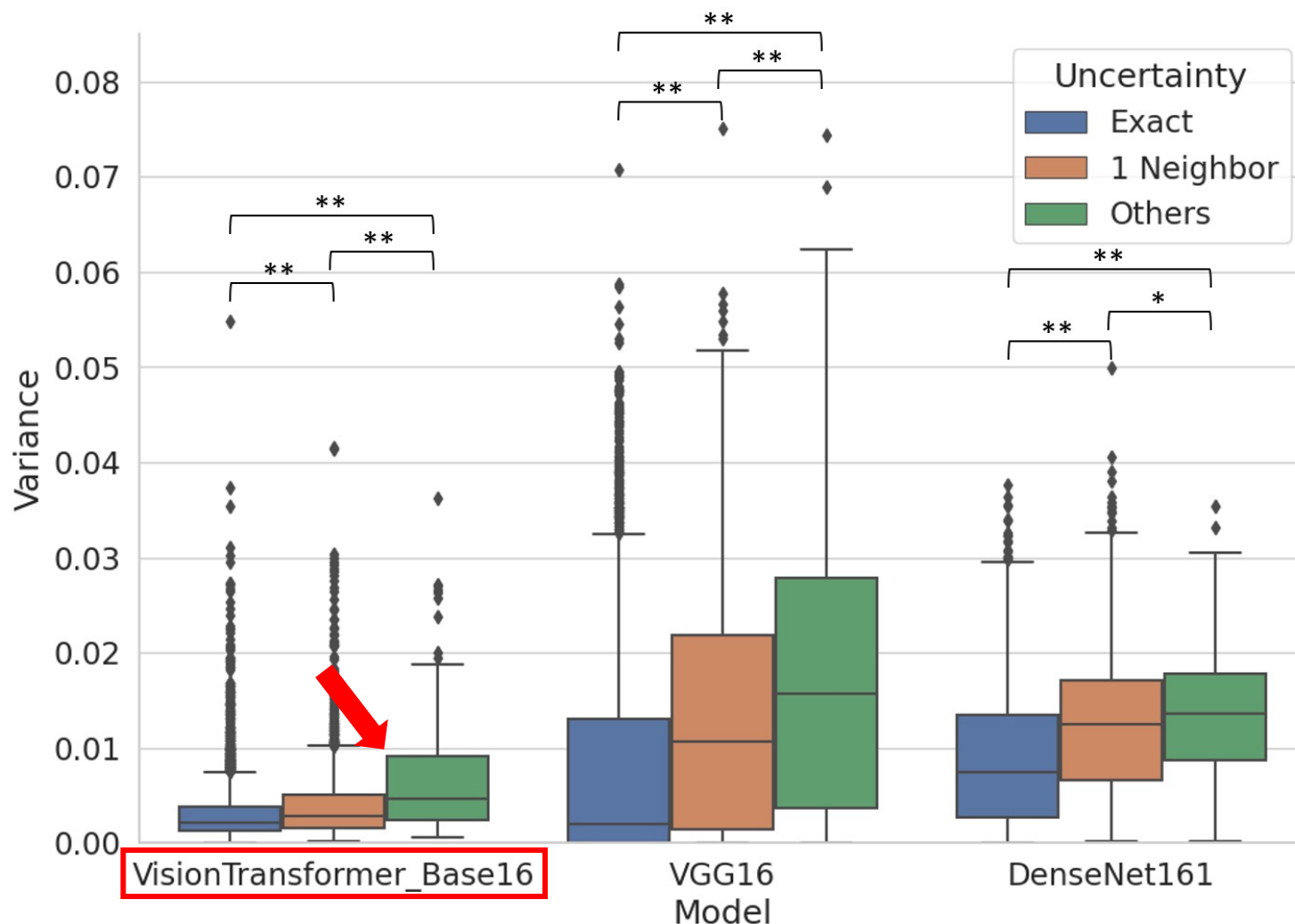
18



Exact : 予測と正解が一致したもの
1 Neighbor : 予測と正解の誤差が1クラスあるもの
Others : 2クラス以上予測を外したもの

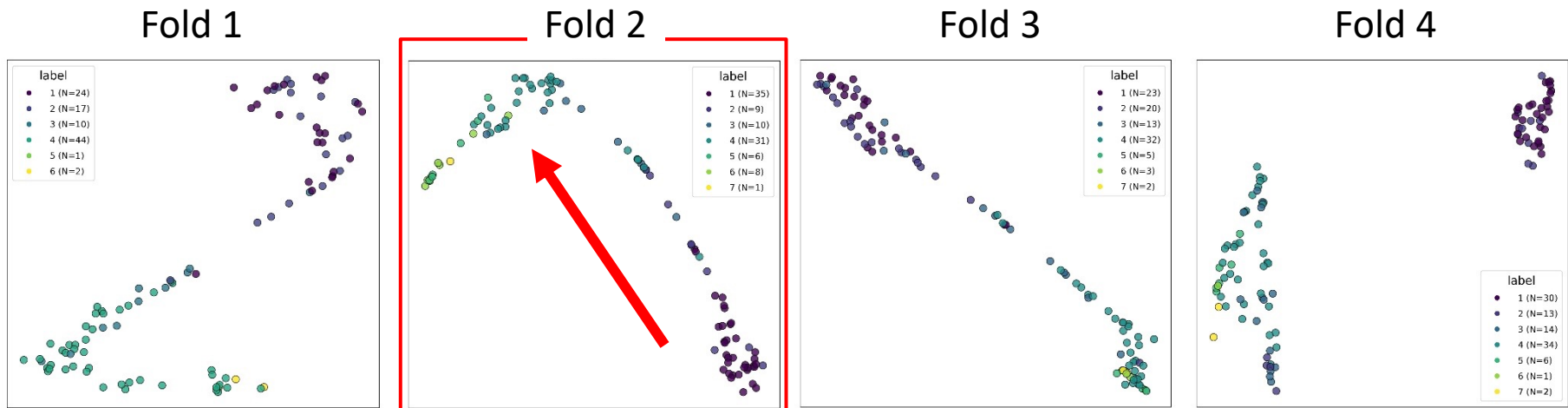
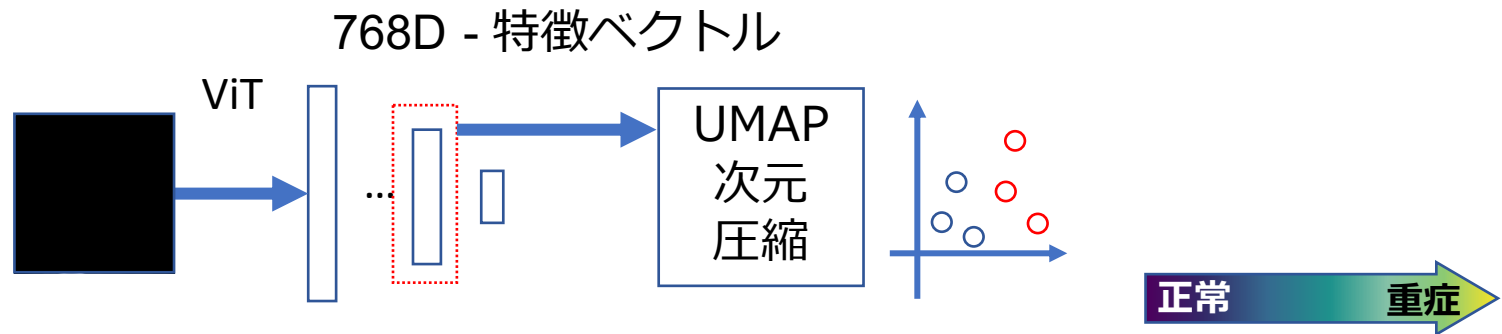
** : $p < .01$
* : $p < .05$

不確実性の評価結果 (Softmax Probabilityの分散) 19



- 分散が大きいことは毎回の予測に確信を持っていないことを示す
- ViTが最も確信を持って予測している(青・オレンジと緑を区別できている)

結果：ViT 特徴ベクトルのUMAP解析 20



Exact class accuracy:

0.673

1-Neighbor Class Accuracy:

0.980

0.650

0.930

0.561

0.929

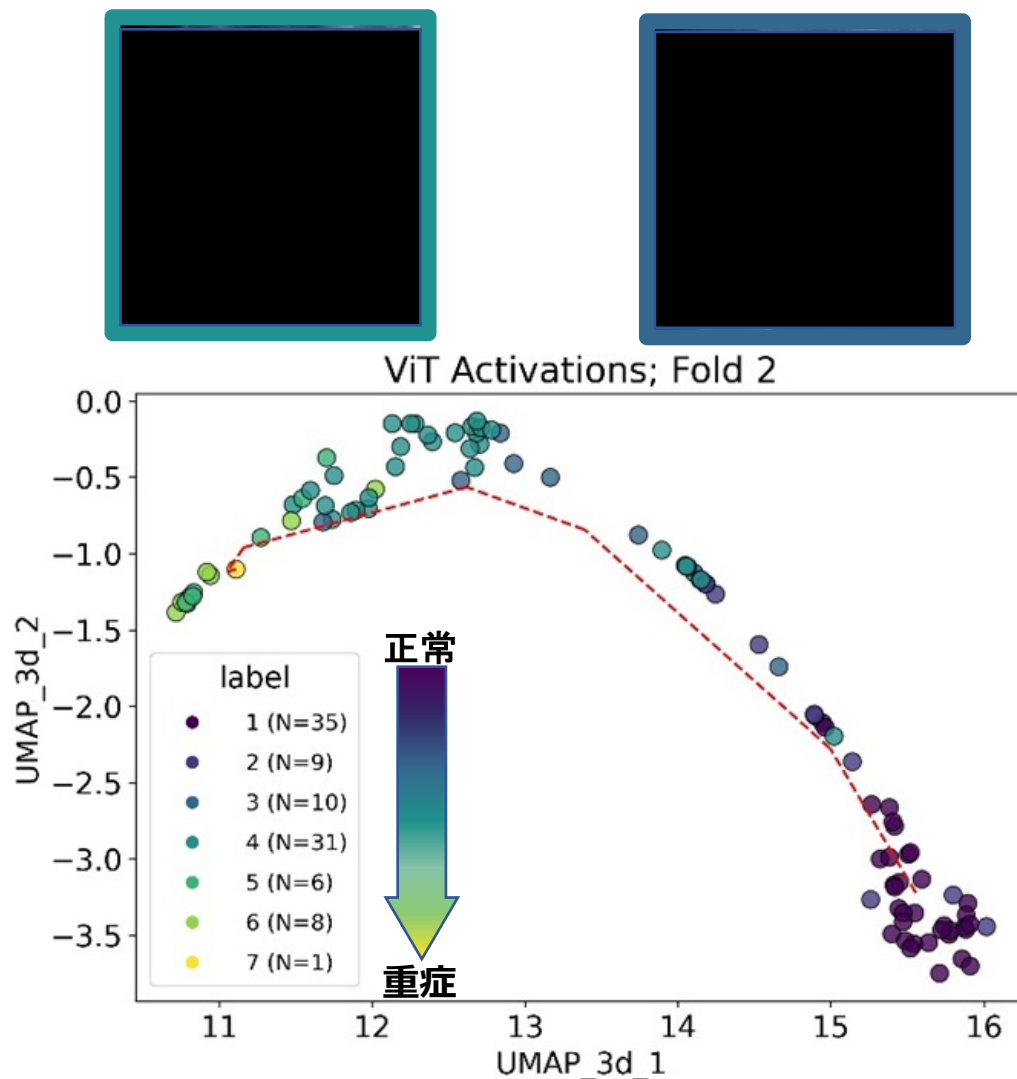
0.650

1.000

結果：ViT UMAP解析 (成功例)

21

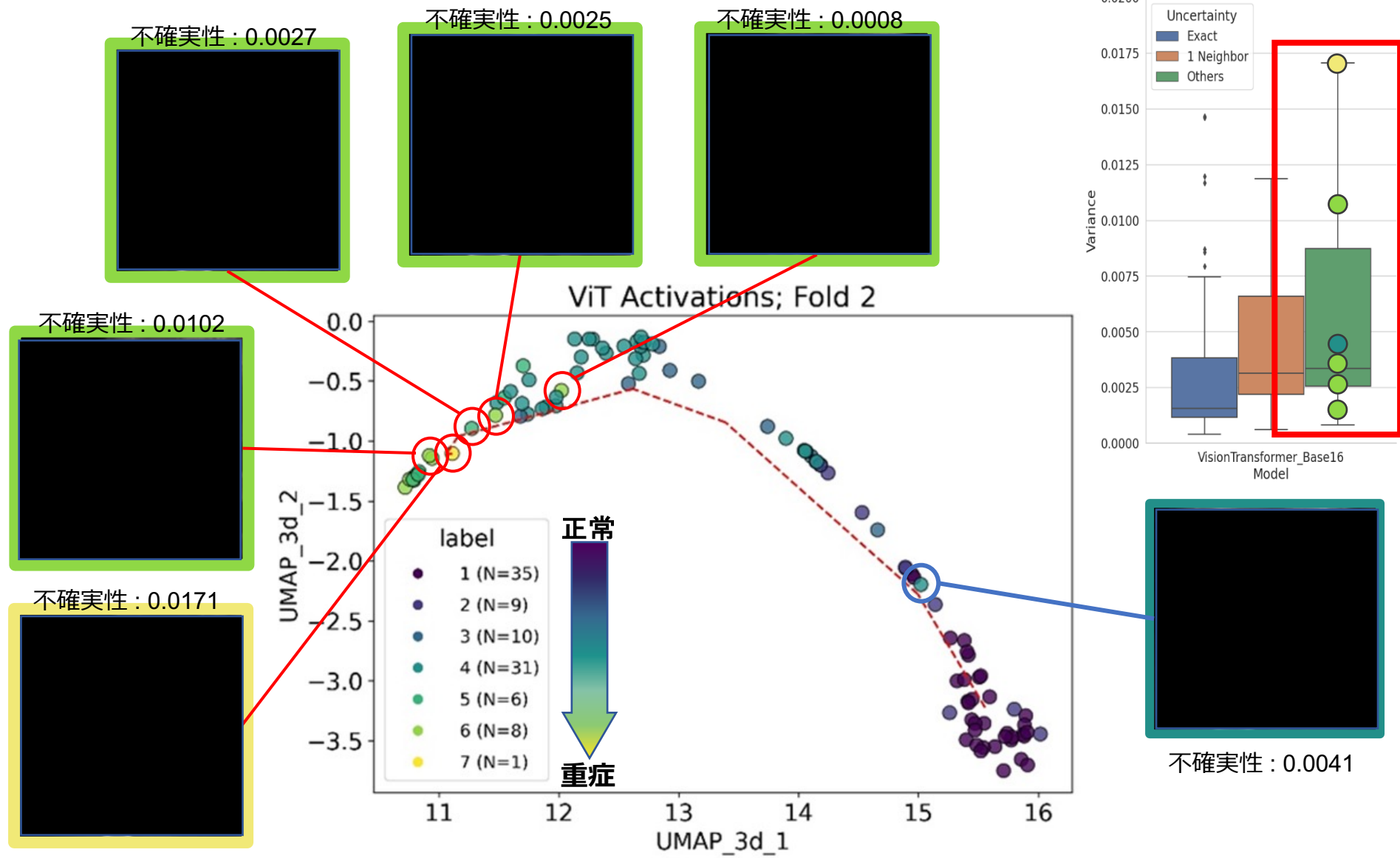
(Exact Class Accuracy : 0.65 ; 1-Neighbor Class Accuracy : 0.93)



結果：ViT UMAP解析 (失敗例)

23

(Exact Class Accuracy : 0.65 ; 1-Neighbor Class Accuracy : 0.93)



結論 & 今後の課題

- 3つの深層学習モデルを用いてHip OAの自動分類を行った
 - 全てのモデルでExact Class Accuracyが60%以上
1-Neighbor Class Accuracyは90%以上の分類精度が得られた
 - 396枚の画像から疾患の進行状況を高精度で予測できた
 - アノテーションが施されていない数千症例の大規模なデータセットに対するアクティブラーニングが可能
 - 予測に確信を持っていない症例を外れ値として、専門医にアノテーションを検討してもらい、学習データに加える。
 - 大規模データベースのアノテーションに携わる医者の負担を減らすことができる
- ➡ 信頼性の高いデータベースの構築

ご清聴ありがとうございました!

Feel free to contact me at
masuda.masachika.mp2@is.naist.jp