

脳画像解析のための 拡張成分分析法

佐賀大学医学部

川口淳

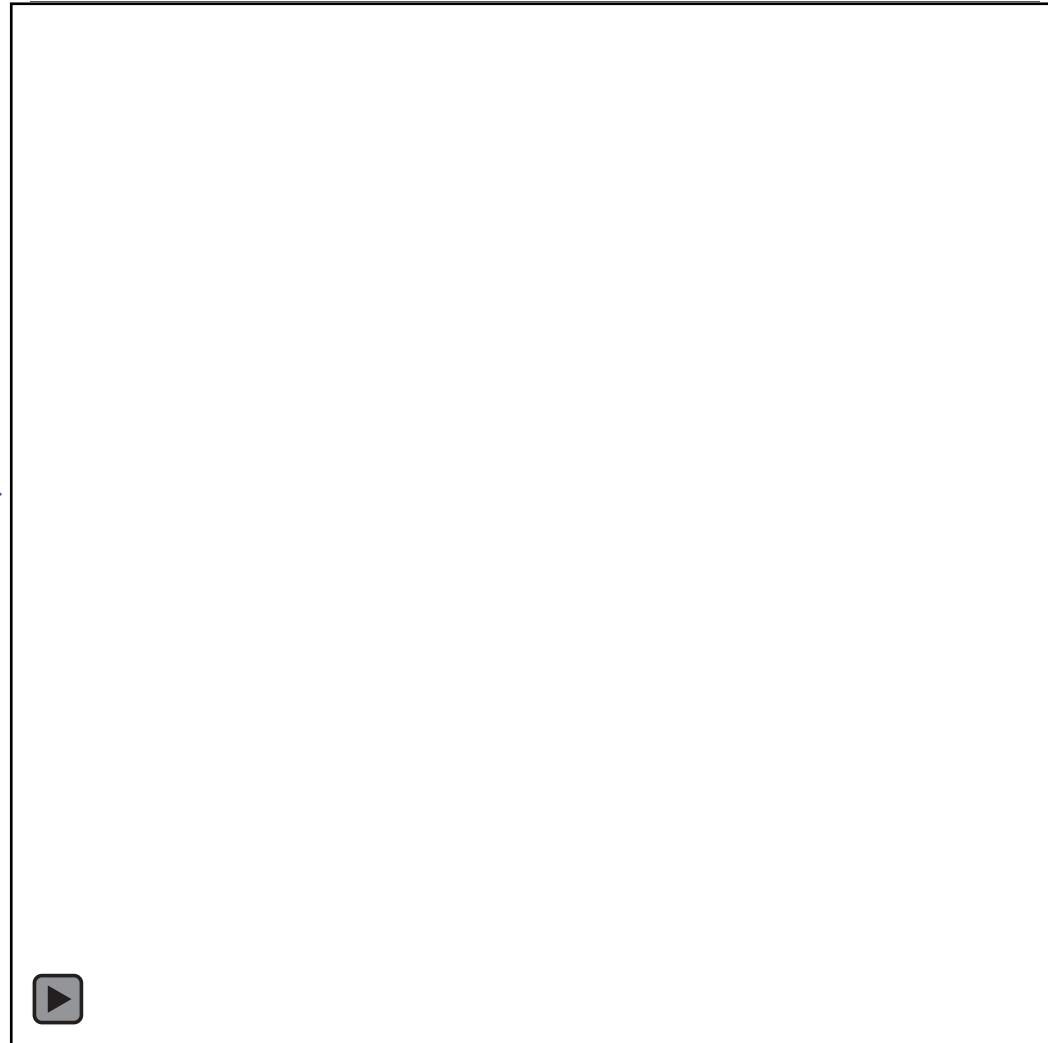
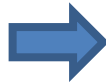
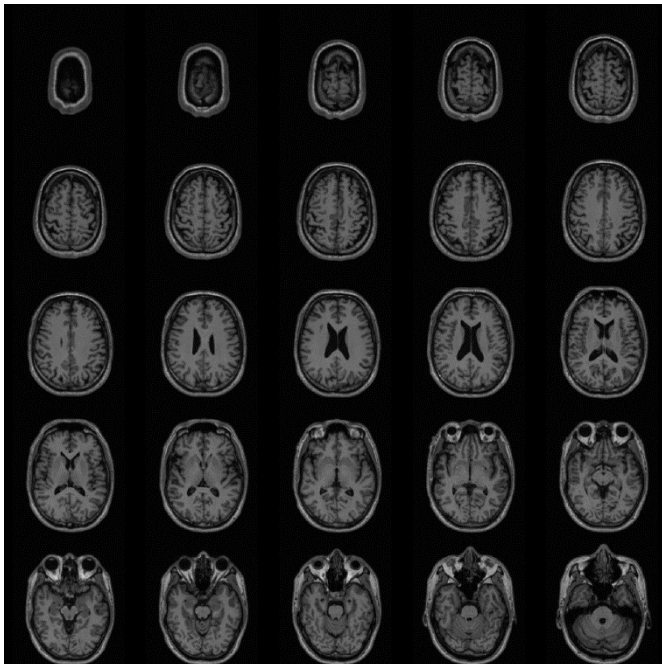
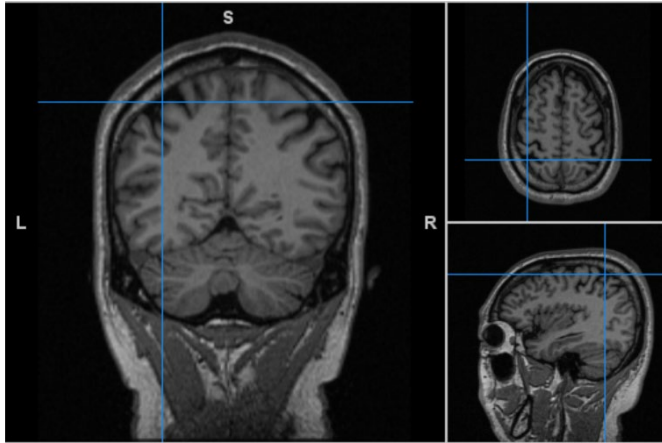
akawa@cc.saga-u.ac.jp

内容

- 脳画像データ解析
- アルツハイマー病における脳画像形態解析
- スコア化のためのマルチ教師付きスパース成分分析
- ハーモナイゼーション

2025年7月23日
東北大学

個人の脳



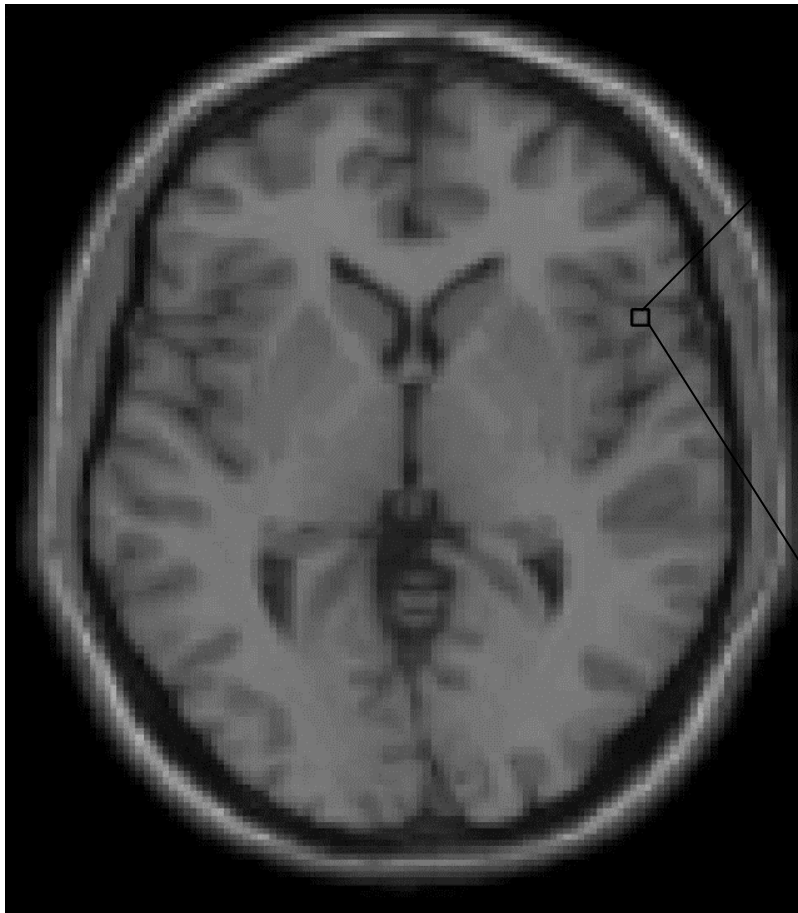
脳画像データ

川口淳 (2018). 画像データの解析.
医学統計学ハンドブック; 714-726

(B) 横断面(1スライス)

(C)色と数値を対応

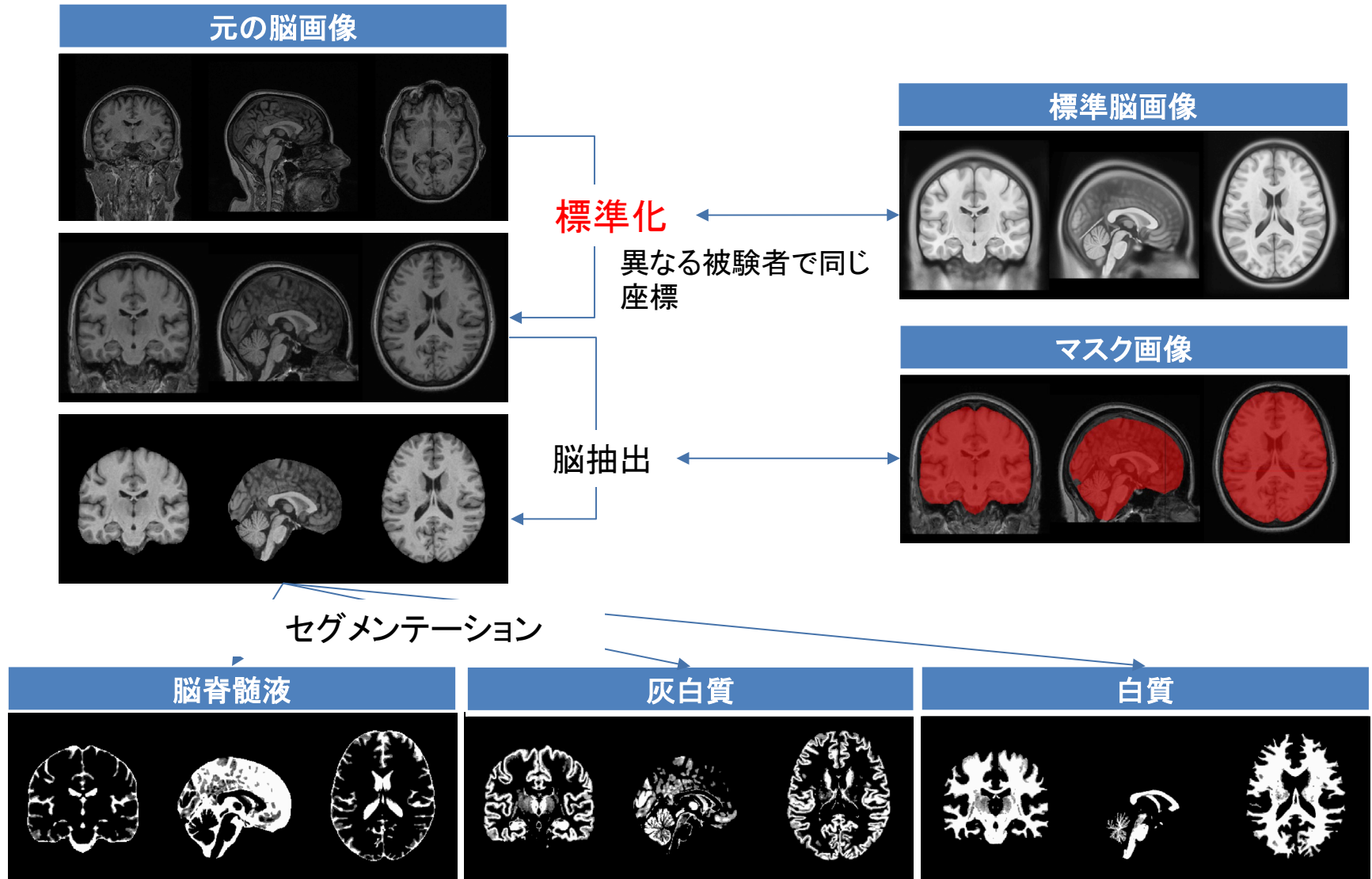
縦



97	95	95	94	94	94	93	90
97	93	93	96	89	93	91	88
97	94	94	97	89	93	92	89
98	93	93	94	83	85	83	79
98	95	95	95	82	83	80	76
97	94	94	94	82	82	81	77
96	93	93	93	78	79	78	
95	96	96	95	80	79	78	

横

前処理



画像データ(単純例)

Aさんの画像

3	1
6	2

Bさんの画像

1	8
5	3

Cさんの画像

8	2
5	6

Dさんの画像

8	7
4	5

ピクセル番地

2	(2,1)	(2,2)
1	(1,1)	(1,2)
	1	2

【群間比較】
検定 + 多重比較補正

群間差あり

(2,1)	(2,2)
(1,1)	(1,2)

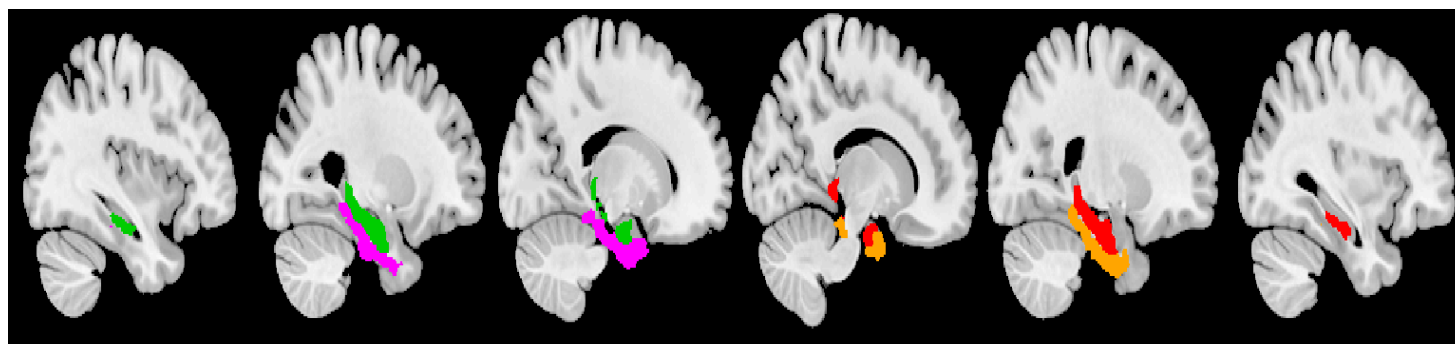
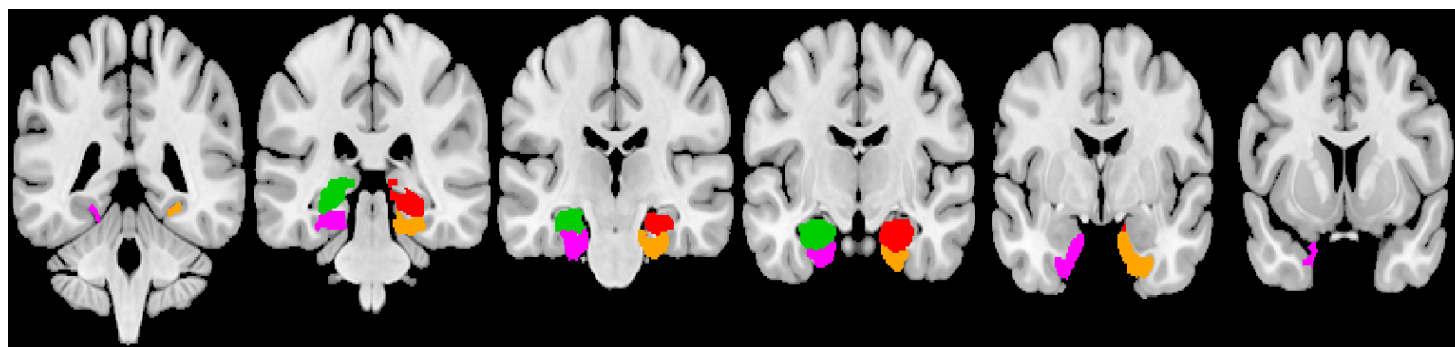
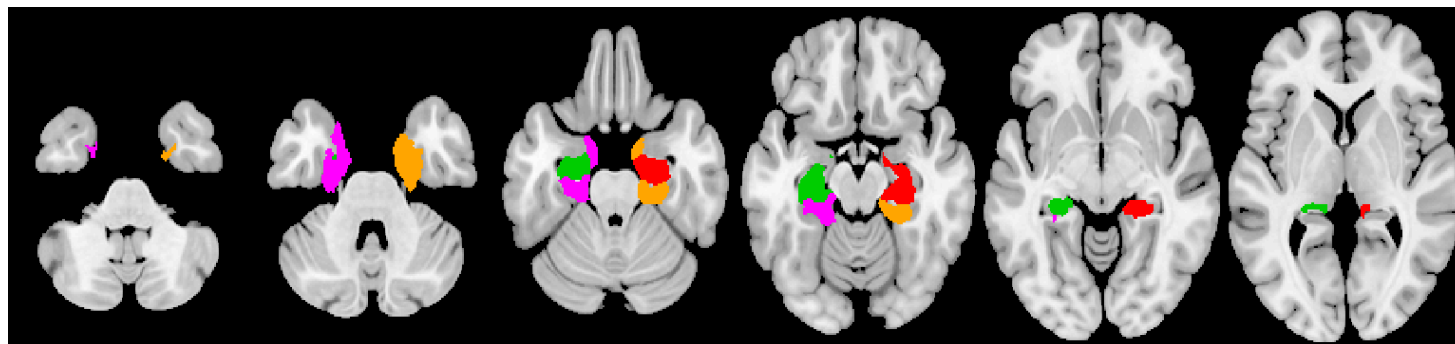
	画像情報 $s(v)$: ピクセル番地 v の輝度				臨床情報 群
	$s((1,1))$	$s((1,2))$	$s((2,1))$	$s((2,2))$	
Aさん	6	2	3	1	健常
Bさん	5	3	1	8	健常
Cさん	5	6	8	2	病気
Dさん	4	5	8	7	病気

【予測】
次元縮小 + 回帰分析

解剖学的领域

左海馬, 右海馬

左海馬傍回, 右海馬傍回

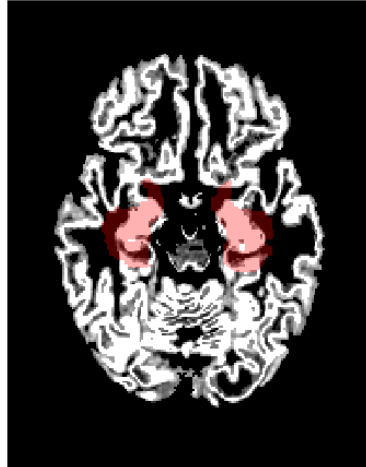


認知症と脳形態

- ▶ アルツハイマー病の初期段階から脳の萎縮が始まる

正常な脳
MMSE=29

OAS30898 MR d0358 925



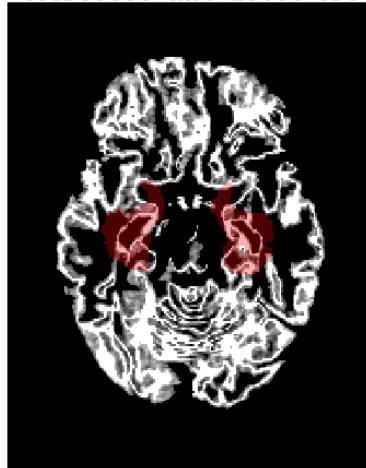
OAS30898 MR d0358 925



76歳
女性

アルツハイマー
病患者の脳
MMSE=19

OAS30999 MR d0153 481



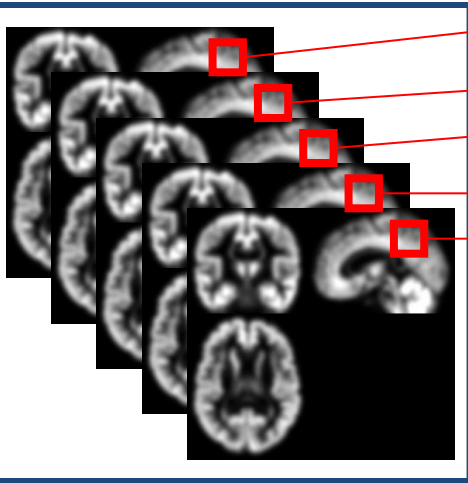
OAS30999 MR d0153 481



massive t検定

Group A

あるボクセル値

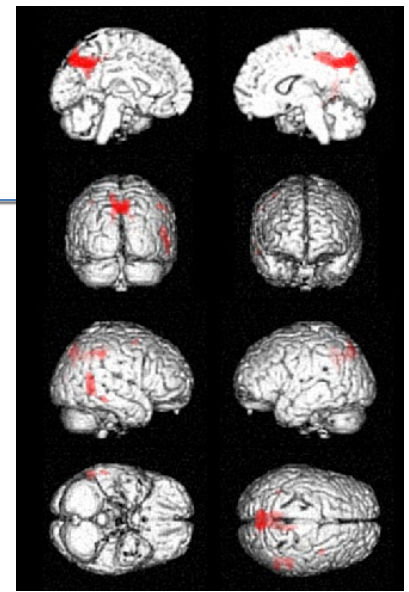


- 50
- 48
- 37
- 45
- 53

平均 \bar{X}

【群間比較】
検定+多重比較補正

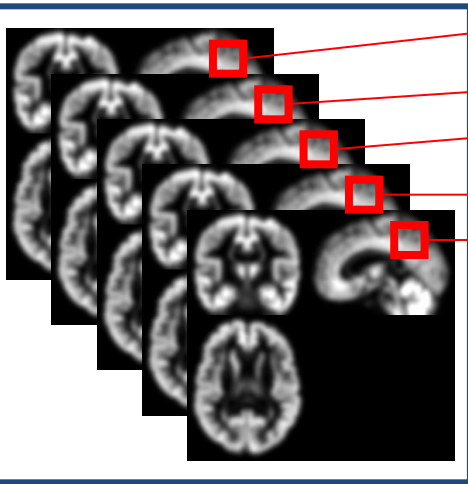
	画像情報 s(v): ピクセル番地vの輝度				群
	s((1,1))	s((1,2))	s((2,1))	s((2,2))	
Aさん	6	2	3	1	健常
Bさん	5	3	1	8	健常
Cさん	5	6	8	2	病気
Dさん	4	5	8	7	病気



$$V = \frac{1}{m+n-2} \left\{ \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{i=1}^n (Y_i - \bar{Y})^2 \right\}$$

全ボクセル対象
のとき,
多重比較補正

Group B



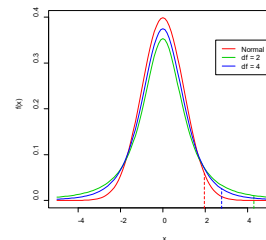
- 47
- 32
- 48
- 45
- 33

平均 \bar{Y}

検定統計量 $T_0 = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)V}}$



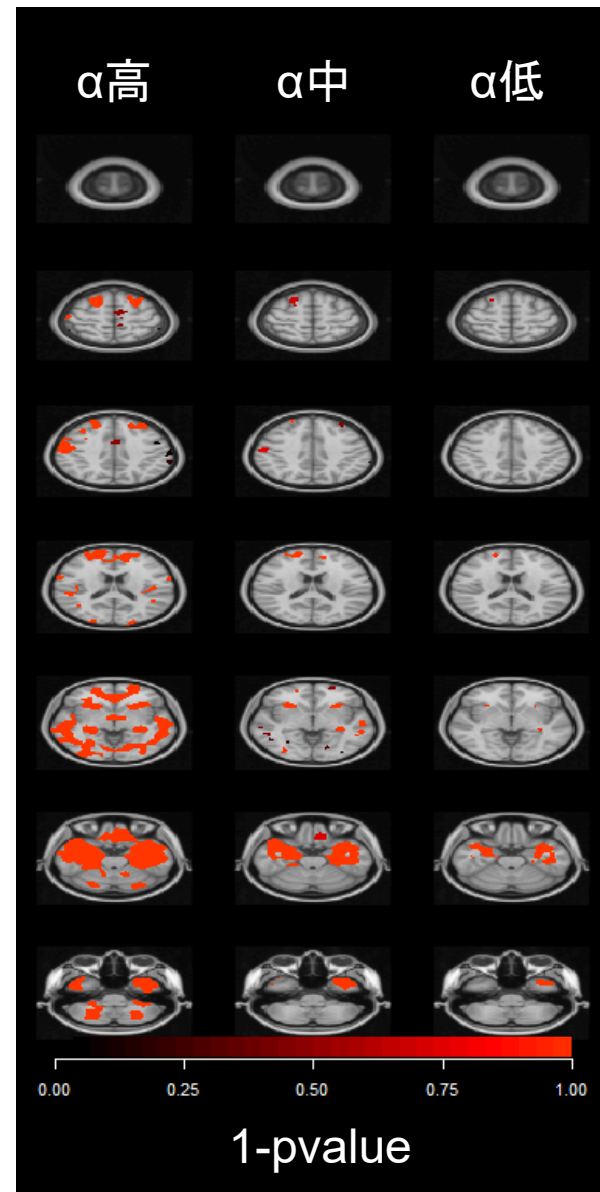
→ $p < 0.???$
有意差あり
なし



解析例

OASISデータ, <http://www.oasis-brains.org/>
 (Open Access Series of Imaging Studies)
 ワシントン大学アルツハイマー病研究グループ

- 国際的評価法「臨床的認知症尺度」
 (CDR)
 - » 正常(=0), 疑い(=0.5), 軽度(=1), 中等度(=2), 高度(=3)
 - » CDR \geq 1をアルツハイマー病(AD) 疑い
 - » CDR = 0を健常(NC)
- 灰白質T1強調画像の品質が良いAD 群
 22名に対して, 年齢, 性別, 脳内容積,
 教育歴をマッチングさせた健常者(NC)群
 22名を選択した.
- Massive t-検定の結果
 - » 有意水準 α で有意な領域を表示

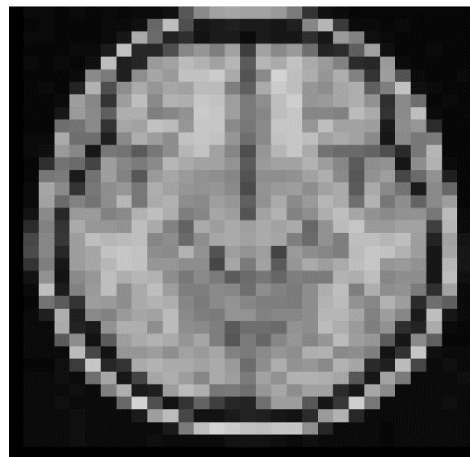


関心領域と個人スコア

- 領域単位にボクセル値を合計≒領域の脳容積
 - ① 解剖学的に定義される領域
 - ② データ駆動型に定義される領域
- ボクセル単位での解析より結果の解釈がしやすい
- 解析における入力変数の数が減らせる(次元縮小)

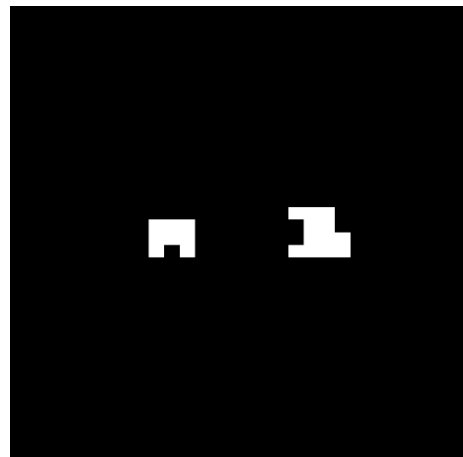
	画像情報 s(v): ピクセル番地vの輝度				臨床情報
	s((1,1))	s((1,2))	s((2,1))	s((2,2))	群
Aさん	6	2	3	1	健常
Bさん	5	3	1	8	健常
Cさん	5	6	8	2	病気
Dさん	4	5	8	7	病気

【予測】
次元縮小+回帰分析



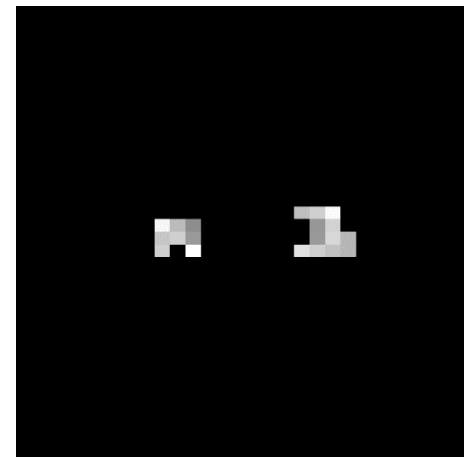
個人の脳

X



関心領域

=



ボクセル値合計 = $3.2 + 3.4 + 2.7 + \dots$
 = 185.9 容積値 (個人スコア)

脳画像のスコア化

元画像

スコア化

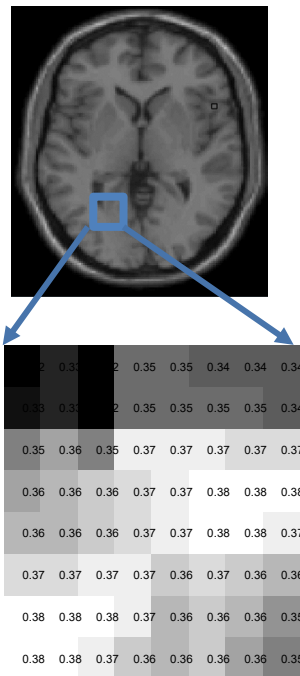
臨床情報との解析

個人の3D脳画像

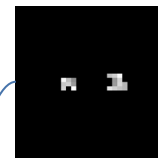
重み(同一)

個人スコア

相関

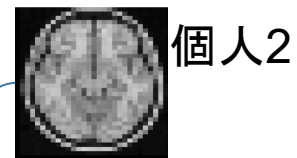


X

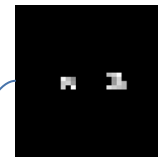


=

185.9 s_1

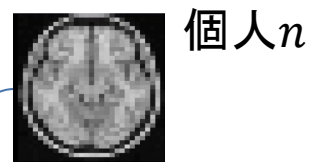


X

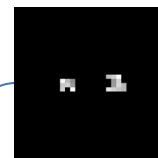


=

162.4 s_2



X

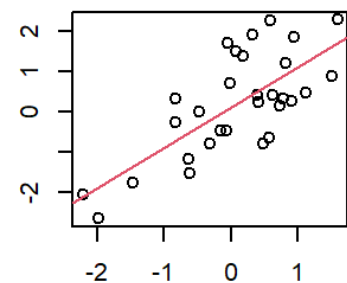


=

193.1 s_n

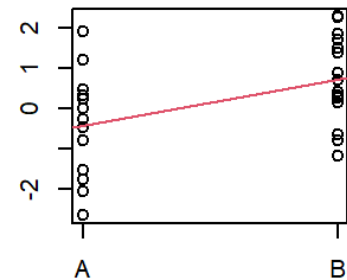
スコア

スコア



臨床情報

群間比較



臨床情報

画素値
(数値データ)

画像

$X_0(n \times N)$

s

重み...

- ①解剖学的ROI(AALなど)
- ②データから推定(主成分分析など)

スコアを求める

主成分分析のアルゴリズム

$$\partial L(\mathbf{w}) / \partial \mathbf{w} = 0$$

$$L(\mathbf{w}) = \mathbf{t}^\top \mathbf{X} \mathbf{w} - \eta \mathbf{w}^\top \mathbf{w} \quad \Rightarrow \quad \frac{\partial L(\mathbf{w})}{\partial \mathbf{w}} = 2\mathbf{t}^\top \mathbf{X} - 2\eta \mathbf{w} \quad \Rightarrow \quad \tilde{\mathbf{w}} = (\mathbf{t}^\top \mathbf{X}) / \eta$$

$$\hat{\mathbf{w}} = \frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|}$$

Deflation step $\tilde{\mathbf{X}} = \mathbf{X} - \hat{\mathbf{X}}$

$$\hat{\mathbf{X}} = \mathbf{s} \mathbf{u}^\top, \quad \mathbf{u} = \mathbf{X}^\top \mathbf{s} / \mathbf{s}^\top \mathbf{s}$$

R 回繰り返す $\mathbf{S} = (\mathbf{s}^{(1)}, \dots, \mathbf{s}^{(R)}), \mathbf{W} = (\mathbf{w}^{(1)}, \dots, \mathbf{w}^{(R)})$

$$\mathbf{S} = \mathbf{X} \mathbf{W}$$

予測例

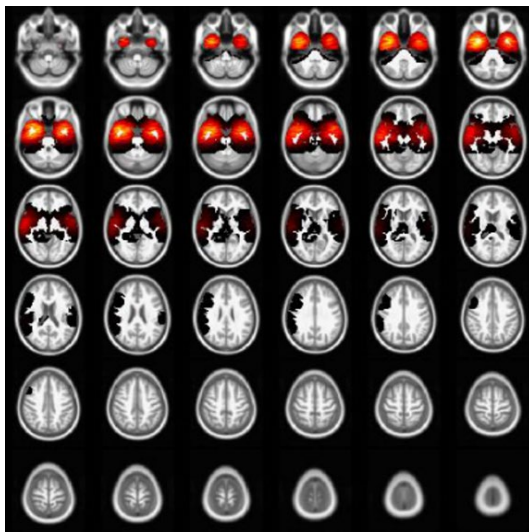
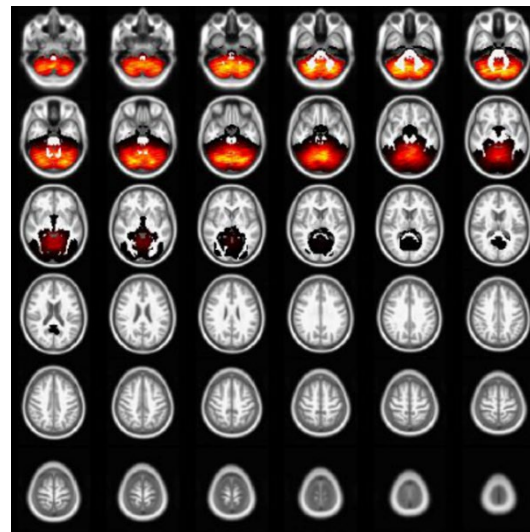
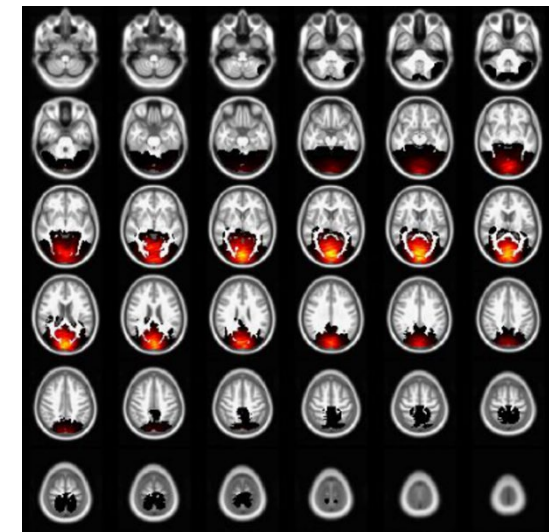
	Dementia	Normal	p
n	52	54	
AGE (mean (sd))	75.41 (7.18)	74.93 (4.89)	0.684
PTGENDER = Male (%)	31 (59.6)	36 (66.7)	0.582

多重ロジスティック回帰分析の結果

$$\text{診断確率 } p(s_1, s_2, s_3) = \{1 + \exp(-(0.046 + 0.014s_1 + 0.018s_2 + 0.043s_3))\}^{-1}$$

$$\Rightarrow p(AK) = 0.044$$

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.046	0.281	0.162	0.8710
s_1	0.014	0.006	2.329	0.0199
s_2	0.018	0.008	2.257	0.0240
s_3	0.043	0.014	3.051	0.0023


 W_1

 W_2

 W_3

スコアの
回帰係数

重みの
復元

ComBat

モデル式

$$y_{ijv} = \alpha_v + X_{ij}^T \beta_v + \gamma_{iv} + \delta_{iv} \varepsilon_{ijv}$$

institute
subject
voxel

- 簡単にするために、施設A 2人, 施設B 2人
- 施設A: $x = 0$ (参照群), 施設B: $x = 1$
- モデル $y = \beta_0 + \beta_1 x + \gamma + \delta \varepsilon$

γ はバイアスの施設間差を変量として特徴づける量
 δ は誤差分散を変量として特徴づける量

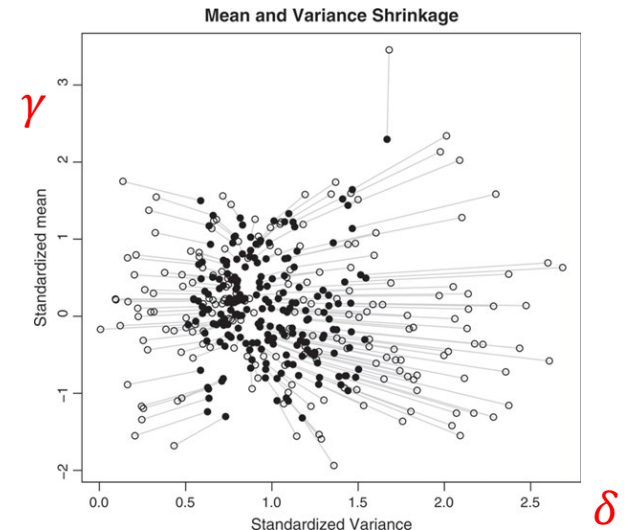
Feature
(ボクセル
など)毎に

$$\begin{pmatrix} \text{施設Aの1人目} \\ \text{施設Aの2人目} \\ \text{施設Bの1人目} \\ \text{施設Bの2人目} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \gamma_1 \\ \gamma_1 \\ \gamma_2 \\ \gamma_2 \end{pmatrix} + \begin{pmatrix} \delta_1 \varepsilon_1 \\ \delta_1 \varepsilon_2 \\ \delta_2 \varepsilon_3 \\ \delta_2 \varepsilon_4 \end{pmatrix}$$

- ベイズ推定による縮小推定
 - » バイアスはボクセル間で大きく異なる必要はない。
 - » 汎化性の向上

combat harmonization

$$y^{combat} = \frac{y - \beta_0 - \beta_1 x - \gamma}{\delta} + \beta_0 + \beta_1 x$$



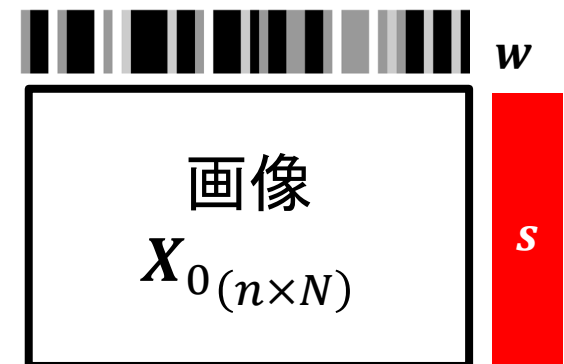
マルチ教師付きスパース成分分析

- 主成分分析＋マルチ教師 による脳画像スコアリング法
 - » 教師付き学習として、必要な変動を学習する。
 - » 反面教師付き学習として、不要な変動を非学習する

- スコア $s = X_0 B w$ の分散のみならず、ベクトル z_1, z_2 との相関も考慮するために、次の関数を制約付き最大化で重み w を求める

$$L(w) = c_0 s^T s + c_1 s^T z_1 - c_2 s^T z_2 - P_\lambda(w)$$

- » スコア $s = X_0 w$ の分散のみならず、 z_1 と z_2 の相関も
- » z_1 が教師：臨床情報(疾患 or 健常)
- » z_2 が反教師，撮像機種(機種A or 機種B)
- » (c_0, c_1, c_2) : 各学習者(教師)のコントラスト
- » $P_\lambda(w) = \lambda \|w\|_1$



解析データ

<https://adni.loni.usc.edu/>

<https://adni.loni.usc.edu/about/>



- 3D T1強調シーケンスで得られる脳の構造MRI画像データ
 - » 撮像機器は2機種
 - > "Siemens Healthineers" と "GE MEDICAL SYSTEMS" とした
 - > MRI内のコイルがメーカーおよび機種間で異なるため、撮像画像の違い出る
 - > 診療上では問題ないが、解析を行う上では注意が必要

- 被験者57人(健常者30人, AD患者27人)

- マッチングされたデータセット
 - » 年齢は70歳前半(70以上75未満)に絞ったうえで、年齢と性別が2群間で分布に差が出ないようにマッチングを行った.

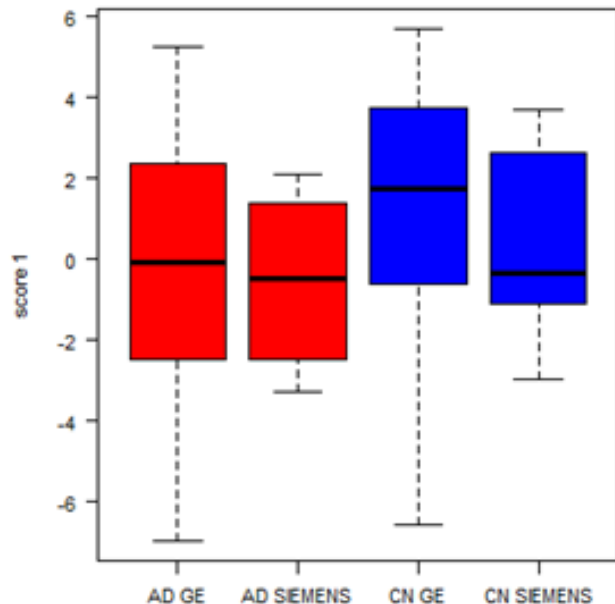
- R packages: msma and mand

結果1 疾患分類と機種によるスコアの分布

$$L(\mathbf{w}) = c_0 \mathbf{s}^\top \mathbf{s} + c_1 \mathbf{s}^\top \mathbf{Z}_1 - c_2 \mathbf{s}^\top \mathbf{Z}_2 - P_\lambda(\mathbf{w})$$

$$(c_0, c_1, c_2) = (1, 0, 0)$$

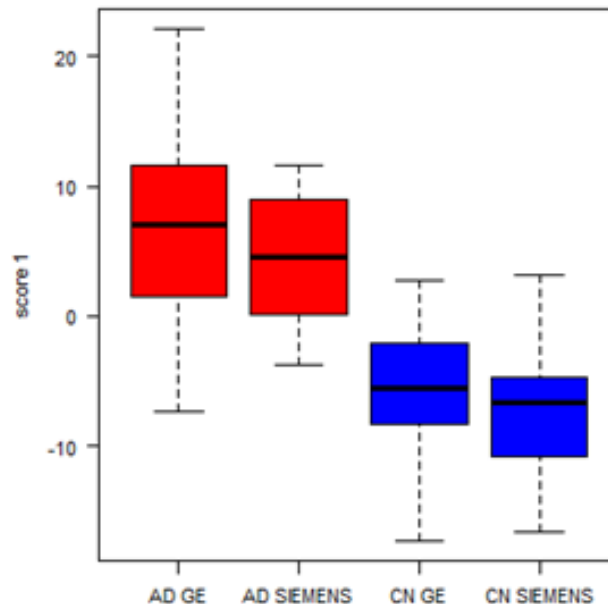
PCA



AUC値=0.623

$$(c_0, c_1, c_2) = (1, 1, 0)$$

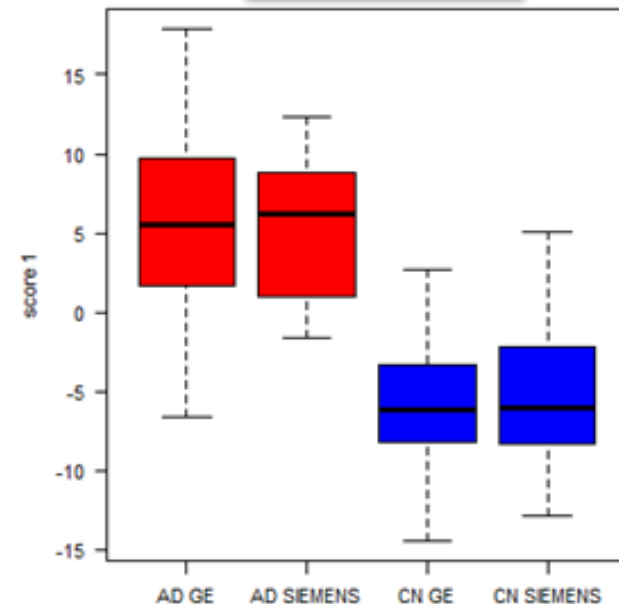
sPCA



AUC値=0.927

$$(c_0, c_1, c_2) = (1, 1, 1)$$

hsPCA

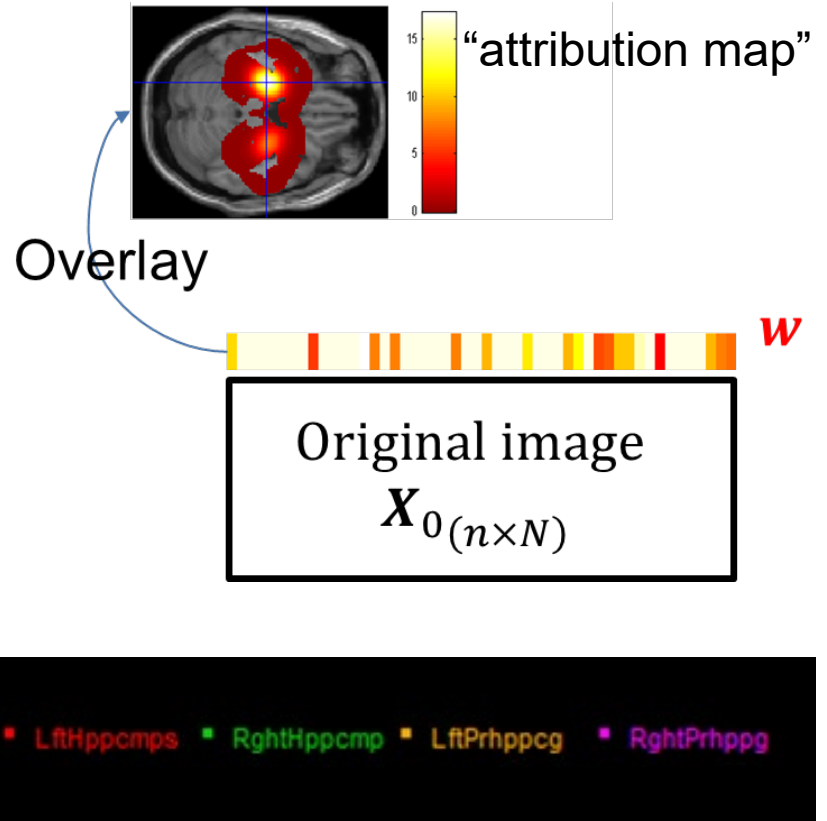
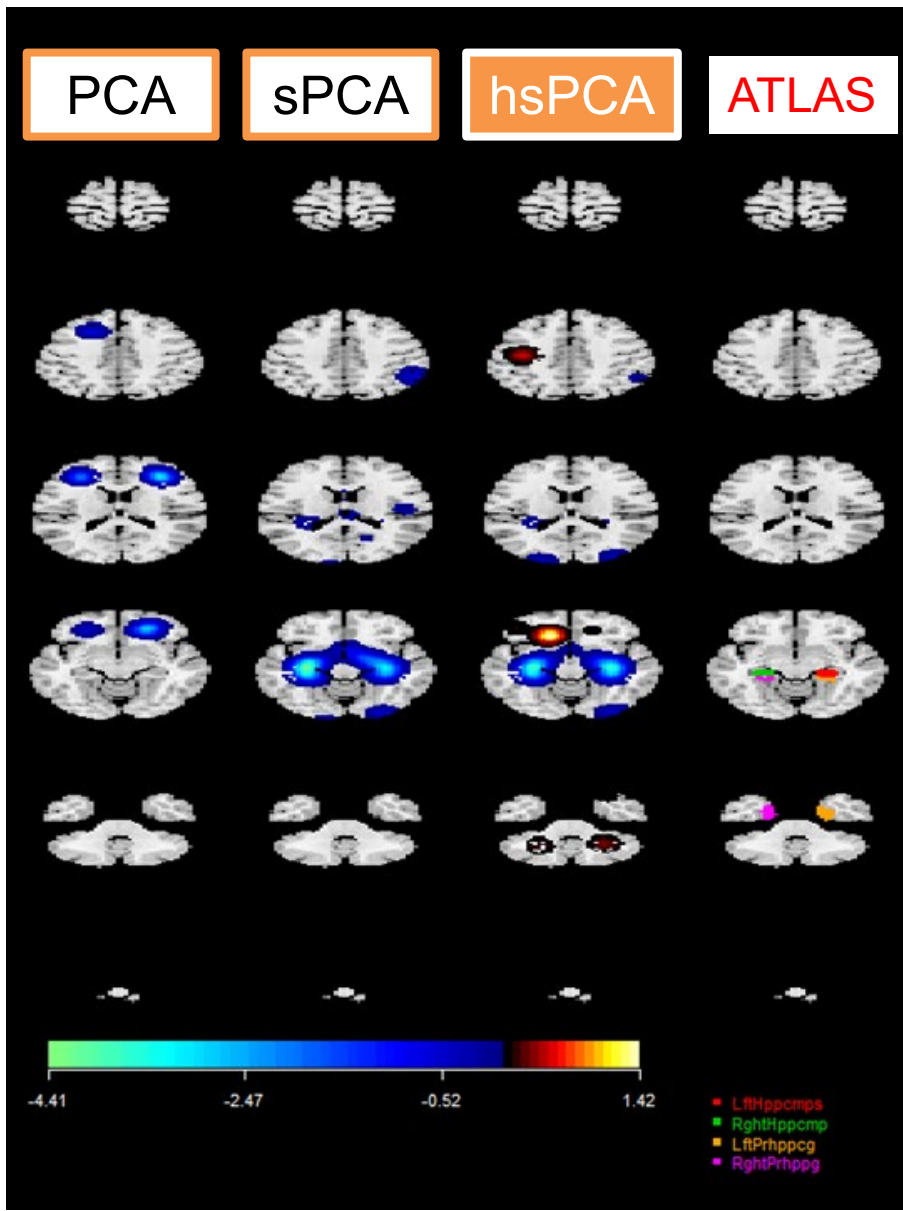


AUC値=0.932

■ AD (Alzheimer's disease), ■ CN (Cognitive Normal)

\mathbf{Z}_2 (左: 機種1, 右: 機種2)

結果2 解剖学的脳領域



提案手法が他の方法より
両側の海馬領域と関連している

説明と精度

トレードオフ

マルチ教師付きスパース成分分析

解析での説明性

線形演算による次元縮小

解釈の説明性

マルチモデル

説明

スコア構造

$$\mathbf{s} = \mathbf{X}\mathbf{w}$$

$$\begin{aligned}\mathbf{s}_X &= \mathbf{X}\mathbf{w}_X \\ \mathbf{s}_Y &= \mathbf{Y}\mathbf{w}_Y\end{aligned}$$

$$\mathbf{s} = \sum_{m=1}^M b_m \mathbf{X}_m \mathbf{w}_m$$

目的関数

$$L(\mathbf{w}) = (1 - \mu)L_0(\mathbf{w}) + \mu L_1(\mathbf{w}) - P_\lambda(\mathbf{w})$$

ただし、 $L_0(\mathbf{w})$ は自己学習項。 $L_1(\mathbf{w})$ は教師学習項。 $P_\lambda(\mathbf{w})$ は正則化項。

タスク

精度

説明

スパース推定

PCA

$$L_0(\mathbf{w}) = \mathbf{s}^\top \mathbf{s}$$

教師付き学習

$$L_1(\mathbf{w}) = \mathbf{z}_1^\top \mathbf{s} \quad L_1(\mathbf{w}) = c_1 \mathbf{z}_1^\top \mathbf{s} + c_2 \mathbf{z}_2^\top \mathbf{s}$$

$$P_\lambda(\mathbf{w}) = \lambda \|\mathbf{w}\|_1$$

PLS

$$L_0(\mathbf{w}) = \mathbf{s}_X^\top \mathbf{s}_Y$$

ハーモナイゼーション

$$L_1(\mathbf{w}) = c_1 \mathbf{z}_1^\top \mathbf{s} - c_2 \mathbf{z}_2^\top \mathbf{s}$$

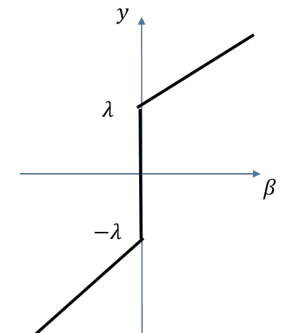
NMF

$$L_0(\mathbf{w}) = \|\mathbf{X} - \mathbf{S}\mathbf{W}\|_F^2$$

一般形

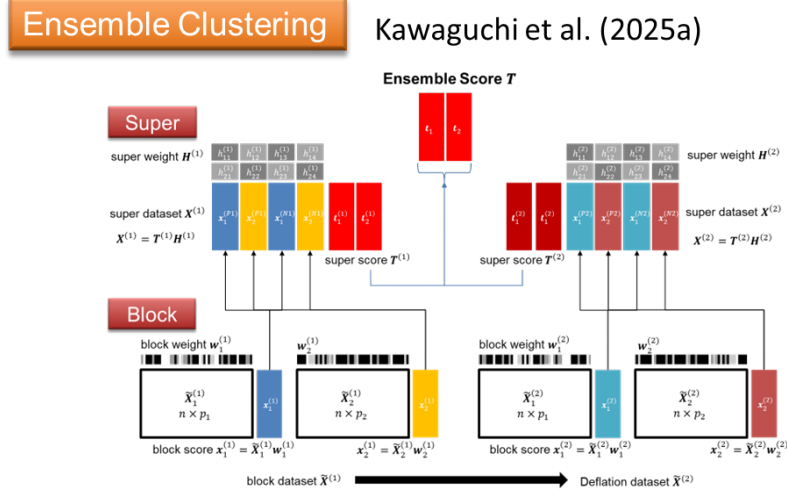
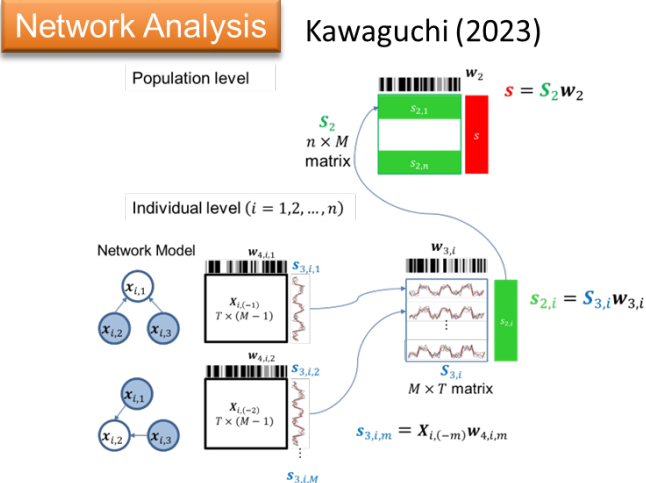
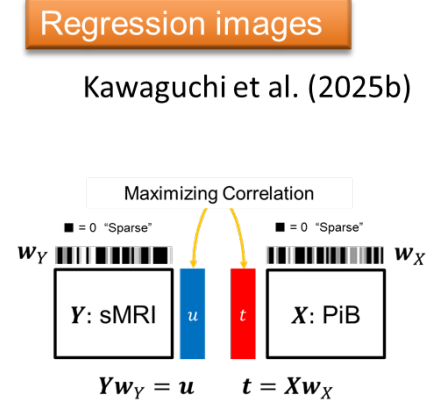
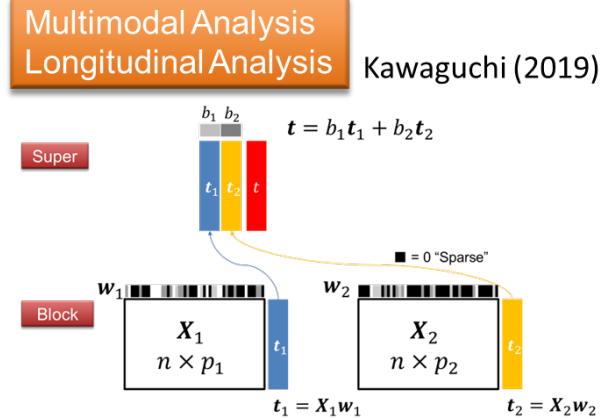
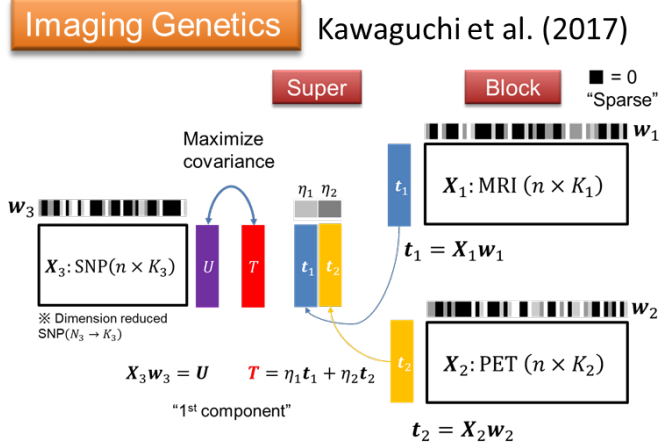
$$L_1(\mathbf{w}) = \sum_{j=1}^k c_j \mathbf{s}^\top \mathbf{z}_j = \mathbf{s}^\top \left(\sum_{j=1}^k c_j \mathbf{z}_j \right) = \mathbf{s}^\top \tilde{\mathbf{z}}$$

各要素は非負



Our Brain Image Scoring Methods

- **Discrimination:** Araki et al. (2013, 2019)
- **Brain imaging and non-image analysis:** Yoshida et al. (2013, 2018)



最適化

目的関数 $L(\mathbf{w})$ を最大にするように重みパラメータ $\mathbf{w} = (w_1, w_2, \dots, w_N)^\top$ を求める。

$$L(\mathbf{w}) = (1 - \mu)L_0(\mathbf{w}) + \mu L_1(\mathbf{w}) - P_\lambda(\mathbf{w})$$

ただし、 $L_0(\mathbf{w})$ は自己学習項。 $L_1(\mathbf{w})$ は教師学習項。 $P_\lambda(\mathbf{w})$ は正則化項。

$L_0(\mathbf{w}) = \mathbf{s}^\top \mathbf{s}$, $L_1(\mathbf{w}) = \mathbf{z}^\top \mathbf{s}$, $P_\lambda(\mathbf{w}) = \lambda \|\mathbf{w}\|_1$ のとき,

$$\begin{aligned} L(\mathbf{w}) &= (1 - \mu)\mathbf{t}^\top \mathbf{X}\mathbf{w} + \mu\mathbf{z}^\top \mathbf{X}\mathbf{w} - \lambda \|\mathbf{w}\|_1 - \eta\mathbf{w}^\top \mathbf{w} \\ &= \{(1 - \mu)\mathbf{t}^\top + \mu\mathbf{z}^\top\}\mathbf{X}\mathbf{w} - \lambda \|\mathbf{w}\|_1 - \eta\mathbf{w}^\top \mathbf{w} \\ &= \{(1 - \mu)\mathbf{t}^\top + \mu\mathbf{z}^\top\} \sum_{j=1}^N \mathbf{x}_j w_j - \lambda \sum_{j=1}^N |w_j| - \eta \sum_{j=1}^N w_j^2 \\ &= \sum_{j=1}^N [\{(1 - \mu)\mathbf{t}^\top + \mu\mathbf{z}^\top\}\mathbf{x}_j w_j - \lambda |w_j| - \eta w_j^2] \end{aligned}$$

最大値を与える w_j ($j = 1, 2, \dots, N$)

$$\tilde{w}_j = h_\lambda(\{(1 - \mu)\mathbf{t}^\top + \mu\mathbf{z}^\top\}\mathbf{x}_j)$$

ただし、 $h_\lambda(y) = \text{sign}(y)(|y| - \lambda)_+$, $(x)_+ = \max(0, x)$

まとめ

- マルチ教師付きスパース成分分析
 - » スコア化による説明, 関連脳領域を示すことによる説明
 - » 学習(教師)と非学習(反面教師)を含め unnecessaryな情報を取り除き, 分類精度向上
- 深層学習などの機械学習法によって画像をハーモナイゼーションする方法も考えられるが, その特徴量も調和されるかは確実ではない (Hu et al. 2023). よってスコアを直接調和する方が好ましいと考える.
- 提案手法で得られるスコアは, 様々な脳画像データに対して調和がとれている上に解剖学的な説明ができ, 信頼性と説明性の高いバイオマーカーと考えられ, 臨床的に有用な解析が可能になると期待される.

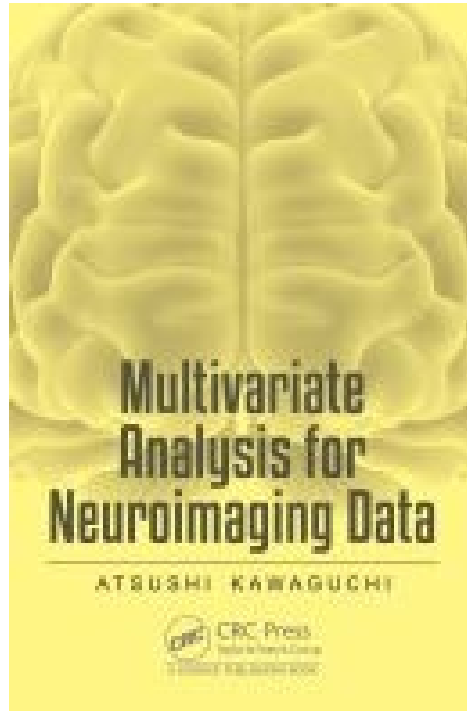
参考文献

- 川口淳 (2013). 脳 MRI データの統計解析. 計量生物学, 33.2: 145-174.
- Kawaguchi, A (2021). Multivariate Analysis for Neuroimaging Data. CRC Press.
- 川口淳 (2022). 脳画像データハーモナイゼーションにおける統計学的解析方法. 日本磁気共鳴医学会雑誌; 42(1) 1-14.
- 川口淳 (2025). 説明性のある脳画像スコアリング法. Medical Imaging Technology; 43(4) 103-109.
- Bayer, Johanna MM, Paul M Thompson, Christopher RK Ching, Mengting Liu, Andrew Chen, Alana C Panzenhagen, Neda Jahanshad, Andre Marquand, Lianne Schmaal, and Philipp G Sämann. (2022). "Site Effects How-to and When: An Overview of Retrospective Techniques to Accommodate Site Effects in Multi-Site Neuroimaging Analyses." *Frontiers in Neurology* 13: 923988.
- Hu, Fengling, Andrew A Chen, Hannah Horng, Vishnu Bashyam, Christos Davatzikos, Aaron Alexander-Bloch, Mingyao Li, et al. (2023). "Image Harmonization: A Review of Statistical and Deep Learning Methods for Removing Batch Effects and Evaluation Metrics for Effective Harmonization." *NeuroImage*, 120125.

Multivariate Analysis for Neuroimaging Data

Atsushi Kawaguchi (著) 215ページ 出版社 : CRC Press

発売日 : 2021/7/1 ISBN-13 : 978-0367255329



Contents

Brain Imaging Data.

Modalities. Format. Preprocess.

Common Statistical Approach.

GLM. Multiple Comparison. Cluster Inference.

Multivariate Approach.

Matrix Decomposition. Sparse Estimation. Dimension Reduction. Multiblock. Classification Method and Prediction model. Machine Learning.

Advance methods.

Multimodal Analysis. Network Analysis. Meta-Analysis.

各方法を解析ソフトRのコードを掲載し解説



mand: Multivariate Analysis for Neuroimaging Data

Several functions can be used to analyze neuroimaging data using multivariate methods based on the 'msma' package. The functions used in the book entitled "Multivariate Analysis for Neuroimaging Data" (2021, ISBN-13: 978-0367255329) are contained. Please also see Kawaguchi et al. (2017) <[doi:10.1093/biostatistics/kxx011](https://doi.org/10.1093/biostatistics/kxx011)> and Kawaguchi (2019) <[doi:10.5772/intechopen.80531](https://doi.org/10.5772/intechopen.80531)>.

Version: 1.0
 Depends: R (≥ 3.5), [msma](#)
 Imports: [oro.nifti](#), [oro.dicom](#), [imager](#), [caret](#)
 Suggests: [knitr](#), [rmarkdown](#)
 Published: 2021-06-29
 Author: Atsushi Kawaguchi [aut, cre]
 Maintainer: Atsushi Kawaguchi <[kawa_a24 at yahoo.co.jp](mailto:kawa_a24@yahoo.co.jp)>
 License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL (≥ 2)]

R package mand

