

令和元年度  
学士学位論文

MRI 脳画像からの年齢推定に向けた  
データ分布補正

Correction of Data Distribution for Age Prediction  
from MRI Brain Images

1200370 元木 拓実

指導教員 吉田 真一

2020年2月28日

高知工科大学 情報学群

# 要旨

## MRI 脳画像からの年齢推定に向けた データ分布補正

元木 拓実

近年、畳み込みニューラルネットワークによる医用画像の分析が活発に行われている。その中でも、MRI によって撮像された脳画像から年齢を推定する研究が多く発表されている。しかし、ある研究ではデータ全体の年齢の平均値で推定してしまう問題が起きており、データの年齢分布の偏りが原因であるとされている。

本研究では、MRI 脳画像を用いた年齢推定においてデータの分布を補正することで推定精度を向上させることが可能であるか検証する。データ分布の補正手法としては、トレーニングデータの年齢層の割合を複製することで補正する。年齢推定は、MRI で撮像された 3 次元構造の灰白質画像を 3 次元畳み込みニューラルネットワーク (3D-CNN) に入力することで推定値を出力する。推定精度の比較は、トレーニングデータを補正しなかった場合と複製により補正した場合とで評価する。評価としては、平均絶対誤差 (MAE) を用い、5 分割の交差検証によって算出される MAE の平均値を最終的な評価とする。

検証実験の結果、補正なしの場合では MAE は 9.63 歳 (標準偏差 6.81) に対して、複製により補正した場合では MAE は 8.96 歳 (標準偏差 6.58) となった。また、推定年齢と実年齢の相関は補正なしでは 0.21、複製による補正ありでは 0.29 となった。このことから、補正ありでは補正なしと比較して推定年齢と実年齢の相関が上がり、幅広い年齢層で推定できるようになった。

キーワード 3 次元畳み込みニューラルネットワーク (CNN), 核磁気共鳴画像 (MRI), 年齢推定, 医用画像

# Abstract

## Correction of Data Distribution for Age Prediction from MRI Brain Images

In recent years, medical image analysis using convolutional neural networks has been actively conducted. Among them, many studies have been published to predict age from brain images obtained by MRI. However, there is a problem that all the predicted age is near the mean age.

In this study, we think the cause of the problem is data distribution and the MRI data age distribution is biased. We propose the correction method of the data distribution. Data distribution is corrected by duplicating the proportion of the age group of the training data. Proposed 3D convolutional neural network (3D-CNN) model inputs the gray matter image of the three-dimensional brain structure captured by MRI. The comparison of the prediction accuracy is evaluated by comparison between corrected and uncorrected data. As the evaluation, the mean absolute error (MAE) is used, and the mean value of the MAE calculated by the 5-fold cross-validation method was used as the final evaluation.

As a result, the MAE is 9.63 years (S.D. 6.81) without correction, and the MAE is 8.96 years (S.D. 6.58) with correction. The correlation between the predicted age and the chronological age is 0.21 without correction and 0.29 with correction by duplication. From this, the correlation between the predicted age and the chronological age increases with the correction compared to the case without the correction, and is predicted in a

wide age group.

***key words*** 3D Convolutional Neural Network (3D-CNN), Magnetic Resonance Imaging (MRI), Age Prediction, Medical Image

# 目次

第 1 章	序論	1
第 2 章	関連研究	2
2.1	脳疾患と脳老化の関連性 . . . . .	2
2.2	CNN による MRI 脳画像からの個人属性の推定 . . . . .	3
2.3	認知症検知指標のための灰白質画像からの年齢推定 . . . . .	3
第 3 章	機械学習手法	5
3.1	3D Convolutional Neural Network (3D-CNN) . . . . .	5
3.1.1	畳み込み層 . . . . .	7
3.1.2	プーリング層 . . . . .	10
3.2	Batch Normalization . . . . .	11
3.3	Dropout . . . . .	12
第 4 章	提案手法と実験内容	14
4.1	データ補正手法 . . . . .	14
4.2	データセット . . . . .	15
4.3	推定モデル . . . . .	16
4.4	評価手法 . . . . .	18
第 5 章	結果	20
第 6 章	考察	22
第 7 章	結論	23
	謝辞	24

目次

参考文献

26

# 目次

3.1	2D-CNN の基本構造 . . . . .	6
3.2	3D-CNN の基本構造 . . . . .	7
3.3	畳み込み層の基本構造 . . . . .	9
3.4	プーリング層 . . . . .	11
3.5	Dropout . . . . .	13
4.1	データセットの年齢分布 . . . . .	15
4.2	年齢推定モデル . . . . .	16
4.3	K-分割交差検証法 . . . . .	19
5.1	補正なしと提案手法の散布図 ○: 補正なし (相関: 0.21) △: 提案手法 (相 関: 0.29) . . . . .	21

# 表目次

4.1	年齢の度数分布表 . . . . .	15
4.2	ネットワーク構成 . . . . .	17
5.1	年齢推定精度 . . . . .	20



# 第 1 章

## 序論

近年、医療分野において畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用いた医用画像の分析が活発に行われている。その中でも MRI 脳画像から推定される年齢と実年齢の差異は、脳の病変が影響していると考えられている [1]。そのため、MRI 脳画像から年齢の推定が可能になることは今後の医療分野における脳の病変の早期発見に役立つと考えられ、多くの年齢推定に関する研究が発表されている。しかし、MRI 脳画像から個人属性を推定する研究 [2] では、年齢推定のタスクにおいてデータ全体の年齢の平均値で推定してしまう問題が起きている。この問題の原因として、使用したデータでは 40 代から 60 代の年齢分布が極端に多いことが挙げられている。一方、別の研究 [3] ではトレーニングデータにおいて、年齢の偏りが少ないデータには画像変換によるデータ複製、年齢の偏りが多いデータにはランダムなデータ除去をすることによってトレーニングデータ全体の年齢の偏りをなくしている。その結果、実年齢と推定年齢の相関係数は 0.85 と高い相関で年齢推定ができています。そこで本研究では、笹谷らの研究 [2] の年齢推定タスクに対して、トレーニングデータの年齢分布をデータの複製によって補正することで年齢推定精度の向上が可能であるか検証する。

第 2 章では、本研究に関連する研究の概要について示す。第 3 章では、本研究で用いる機械学習手法の概要について示す。第 4 章では、データの分布補正手法、検証実験で使用したデータセット、年齢推定のための 3D-CNN モデル、年齢推定モデルの精度を決定する評価手法について示す。第 5 章では、データを補正しない場合と提案手法により補正した場合の年齢推定結果、またそれぞれにおける推定値と正解値の散布図を示す。第 6 章では、結果からの考察を述べる。第 7 章では、本研究全体のまとめを結論として示す。

## 第 2 章

# 関連研究

本研究では、3次元構造のMRI脳画像を用いて年齢推定を行う。そこで本章では、3次元構造のMRI脳画像を用いた年齢推定に関連する研究について述べる。

### 2.1 脳疾患と脳老化の関連性

脳疾患は、世界中の罹患率や障害の主な原因の一つであり、脳の年齢的な変化が脳疾患に関連があると示唆されている。また、機械学習によってMRI脳画像から年齢を推定できるとされている。このような背景から Kaufmann ら [1] は、機械学習を用いたMRI脳画像からの年齢推定によって脳の老化と脳疾患の関連性を研究している。この研究では、入力画像としてT1強調画像を前処理したものを使用している。データセットは、研究に参加した3歳から96歳の被験者45,615人のMRI脳画像が使用された。機械学習手法としては、Rのxgboostパッケージが用いられた。また、研究対象として挙げられた脳疾患は、自閉症スペクトラム障害、注意欠陥多動性障害、前駆期統合失調症または精神病発症危険状態、統合失調症、精神病スペクトラムの混合診断を持つグループ、双極スペクトラム障害、多発性硬化症、大うつ病性障害、軽度認知障害、アルツハイマー病を含む認知症である。年齢推定は男女それぞれで行い、男性モデルでは相関が0.94、女性モデルでは相関が0.93と性別関係なく高い相関で年齢推定ができていた。また、脳の老化と脳疾患の関連性としては、統合失調症、精神病スペクトラムの混合診断を持つグループ、双極スペクトラム障害、多発性硬化症、大うつ病性障害、軽度認知障害、アルツハイマー病を含む認知症で関連していることが示された。

### 2.2 CNN による MRI 脳画像からの個人属性の推定

近年、MRI で撮像された脳の構造画像からアルツハイマー型認知症といった脳疾患との関連性を導くことができるとされていることから、医用画像処理による診断の自動化や健康状態の予測が望まれている。このような背景から笹谷ら [2] は、MRI 脳構造画像から機械学習を用いて人間の最も基本的な情報である性別と年齢を予測することができるか研究している。また、日常行動から喫煙情報を導くことができるかということも研究している。この研究では、入力画像として T1 強調画像を前処理し、標準脳に合わせられた頭部画像と灰白質画像を使用している。機械学習手法としては、2 クラス分類の性別推定と回帰推定タスクの年齢推定では 2D-CNN と 3D-CNN、3 クラス分類の喫煙量推定では 2D-CNN を用いている。研究結果として、性別推定では 2D-CNN モデルの場合、頭部画像で 97.00%、灰白質画像で 85.32% という精度の推定結果が得られた。また 3D-CNN モデルの場合、頭部画像で 97.01%、灰白質画像で 83.71% という精度の推定結果が得られた。次に年齢推定では、頭部画像を用いた 2D-CNN モデルと 3D-CNN モデルでともに、平均二乗誤差 (MSE) で 9.09 歳、平均絶対誤差 (MAE) で 7.08 歳の誤差で推定する結果が得られた。喫煙量推定では、頭部画像を用いた 2D-CNN モデルで“喫煙中”、“禁煙中”、“未経験”の 3 クラス分類を行い、54.34% という結果が得られた。年齢推定では、データの年齢の平均値を予測値として出力してしまう問題が起きており、原因として使用したデータのうち 40 代から 60 代の年齢分布が極端に多かったことが挙げられている。

### 2.3 認知症検知指標のための灰白質画像からの年齢推定

MRI 脳画像を用いた年齢推定において、推定される年齢と実年齢の差異は初期段階の脳神経変性を発見するための指標として役立つ可能性があると考えられており、近年の研究では脳の老化と統合失調症などの様々な疾患との関係が示されている。このような背景から Wang ら [3] は、MRI 脳画像からセグメンテーションされた灰白質画像を使用して、灰白質から予測された年齢の差異と認知症との関係を研究している。この研究では、入力画像

## 2.3 認知症検知指標のための灰白質画像からの年齢推定

として T1 強調画像を前処理し、標準脳に合わせられた灰白質画像を使用している。データセットは、健康な男女 5,497 人の MRI 脳画像がロッテルダム研究のデータから使用された。機械学習手法としては回帰推定タスクの 10 層の 3D-CNN モデルを用いている。また、トレーニング時にトレーニングデータに対し、年齢の割合が少ないデータには画像変換（シフト、ズームイン/アウト、回転、反転、歪曲、色相陰影）によるデータ複製、年齢の割合が多いデータにはランダムな除去をすることで、トレーニングデータの年齢範囲全体でデータを均等に分散し、学習時のトレーニングデータへの過剰適合を回避している。研究結果として、平均絶対誤差（MAE）が 4.45 歳（標準偏差 3.59）の誤差で推定する結果が得られた。また、相関は 0.85 となっている。さらに、灰白質画像から推定される年齢と実年齢の差異は認知症発生率と有意に関連していることが示された。

このような関連研究から本研究では、笹谷らの研究 [2] の年齢推定タスクに対して、年齢の割合が少ないデータを複製によって補正することで推定精度の向上が可能であるか検証する。また、灰白質画像に年齢に関する情報があると考えられるため、入力データとして灰白質画像を使用する。

## 第 3 章

# 機械学習手法

本章では、本研究で用いる機械学習手法である 3D Convolutional Neural Network (3次元畳み込みニューラルネットワーク) の概要を述べる。

### 3.1 3D Convolutional Neural Network (3D-CNN)

Convolutional Neural Network (畳み込みニューラルネットワーク) は、画像認識や音声認識といった様々な認識タスクにおいて、高い精度の識別性能を持つニューラルネットワークである。CNN のルーツとなったニューラルネットワークの一つに福島らが提案した Neocognitron があり、生物の脳の視覚神経系の構造を参考にしてモデル化している [4]。CNN は、この Neocognitron に対して誤差逆伝播法に基づいた勾配降下法を適用して教師あり学習に改良したもので、LeCun らによって提案された [5]。現在、多くの研究で用いられている CNN は 2D-CNN とも呼ばれ、基本的に 2 次元データを対象としたネットワーク構造となっている。そこで、2 次元情報に奥行き情報を加えて 3 次元データを処理することが可能な 3D-CNN モデルが提案された [6]。本研究では、データセットとして 3 次元データである MRI 構造画像を用いているため、3D-CNN によって脳画像の情報を維持したまま特徴抽出が可能であると考えられる。

2D-CNN の基本構造は図 3.1 のように入力層、畳み込み層 (Convolution Layer)、プーリング層 (Pooling Layer)、全結合層 (Fully Connected Layer)、出力層といった層から構成されている。CNN の特徴としては、畳み込み層とプーリング層を交互に接続した構造を持つことが挙げられている [7]。この畳み込み層とプーリング層により、画像などのデー

### 3.1 3D Convolutional Neural Network (3D-CNN)

タからの局所特徴を抽出することができ、抽出された特徴から全結合層は最終的な識別をする。また図 3.2 より、2D-CNN と 3D-CNN の間には大きな構造の違いはなく、3D-CNN でも畳み込み層とプーリング層によって特徴量を抽出し、全結合層で最終的な識別をしている。異なる点としては、特徴抽出に奥行き情報を得るための軸が加えられており、これによって 3 次元データの特徴抽出を可能としている。

次に、3D-CNN において畳み込み層とプーリング層で行われる処理について説明する。

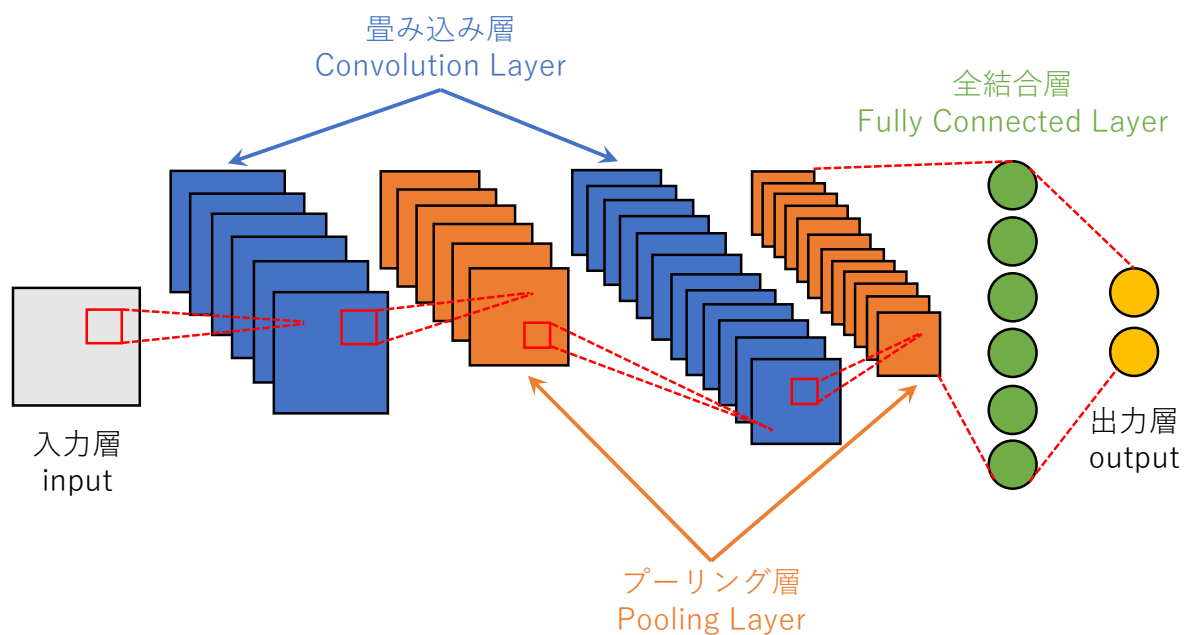


図 3.1: 2D-CNN の基本構造

### 3.1 3D Convolutional Neural Network (3D-CNN)

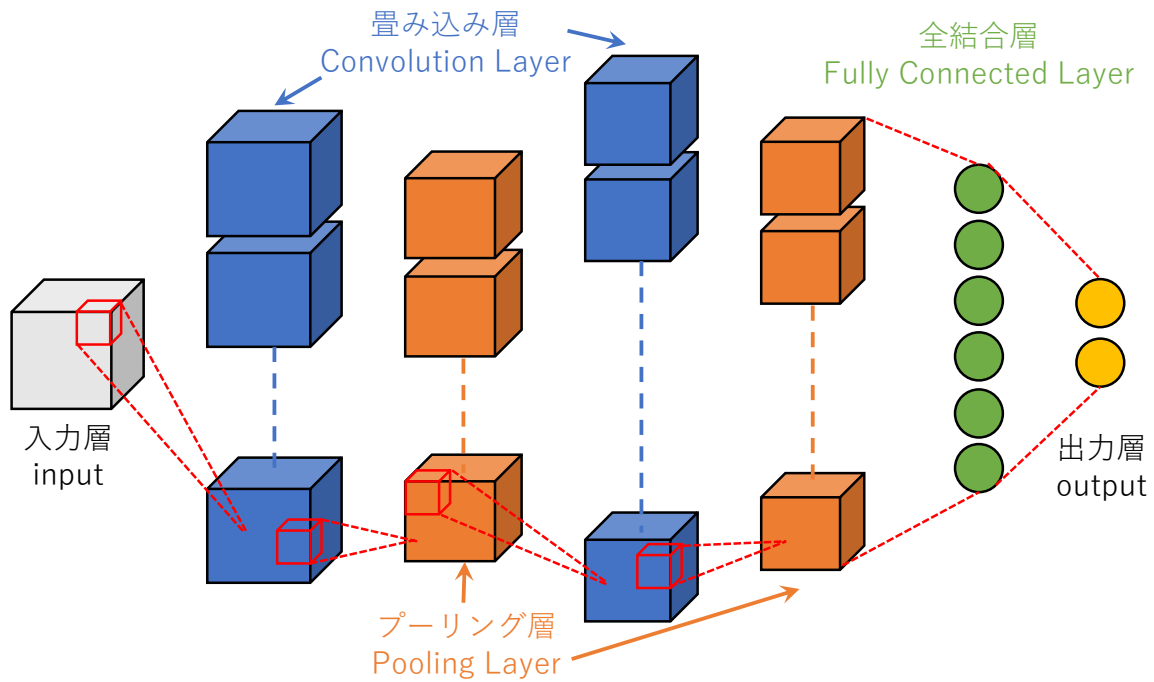


図 3.2: 3D-CNN の基本構造

#### 3.1.1 畳み込み層

畳み込み層の基本構造は図 3.3 のようになっており，図 3.3 は 1 つのフィルタに関する畳み込み層での畳み込み処理を示している．入力画像  $\boldsymbol{x} = (x_{ijkl})$  は横方向，縦方向，奥行き  
のサイズが  $S_x \times S_y \times S_z$  画素の  $N$  チャンルの画像の形をとる．チャンネル数は，入力画像が  
グレースケールするとき  $N = 1$ ，カラーでは RGB の 3 枚で  $N = 3$  になる． $x_{ijkl}$  は式 (3.1)  
で表される．

$$x_{ijkl} = \{(i + i', j + j', l + l', k + k') \mid i' = 0, \dots, S_x - 1, j' = 0, \dots, S_y - 1, \\ l' = 0, \dots, S_z - 1, k' = 0, \dots, N - 1\}. \quad (3.1)$$

畳み込み層では，この入力画像  $\boldsymbol{x}$  に対してフィルタを畳み込む演算を行う．

具体的な計算内容としては，入力画像  $\boldsymbol{x}$  のサイズ  $S_x \times S_y \times S_z$  画素の各チャンネルごとに  
 $L_x \times L_y \times L_z$  のサイズの 3 次元フィルタを畳み込み，その結果を全  $N$  チャンネルにわたって加  
算する．この計算結果は 1 チャンルの画像  $\boldsymbol{u} = (u_{ijl})$  をとる． $\boldsymbol{u}$  はフィルタを  $\boldsymbol{w} = (w_{ijkl})$

### 3.1 3D Convolutional Neural Network (3D-CNN)

としたとき、式 (3.2) のように計算される。ただし、 $P_{ijl}$  は入力画像中の画素  $(i, j, l)$  を始点とするサイズ  $L_x \times L_y \times L_z$  画素のフィルタ領域であり、式 (3.3) で表される。また、式 (3.2) の最後にある  $b_k$  はバイアスであり、各チャンネル  $k$  ごとに全出力ノード間で同一とすることが多い。

$$u_{ijl} = \sum_{k=1}^N \left[ \left[ \sum_{(p, q, r) \in P_{ijl}} x_{pqrk} w_{p-i, q-j, r-l, k} \right] + b_k \right], \quad (3.2)$$

$$P_{ijl} = \{(i + i', j + j', l + l') \mid i' = 0, \dots, L_x - 1, j' = 0, \dots, L_y - 1, l' = 0, \dots, L_z - 1\}. \quad (3.3)$$

フィルタを全画素ではなく、数画素間隔で適用することもあり、その場合はある画素数  $s$  について式 (3.2) の  $P_{ijl}$  を式 (3.4) とし、 $w_{p-i, q-j, r-l, k}$  を  $w_{p-si, q-sj, r-sl, k}$  と置き換えて  $\mathbf{u}$  を計算する。このとき、出力層のノード数は入力層の  $(1/s)^2$  倍になる。この画素間隔  $s$  はストライドと呼ばれ、入力画像サイズが大きいときにネットワークのサイズが大きくなりすぎるのを避けるため、2 以上の  $s$  を選ぶことがある。例として  $s = 2$  にした場合、フィルタは 2 画素ずつ縦横に移動するため、16 個のノードが入力されたときに出力されるノードは 4 個になる。ただし、ストライドを増やすと特徴を取りこぼすことに繋がり、一般に性能は低下する傾向にある。

$$P_{ijl} = \{(si + i', sj + j', sl + l') \mid i' = 0, \dots, L_x - 1, j' = 0, \dots, L_y - 1, l' = 0, \dots, L_z - 1\}. \quad (3.4)$$

このように計算された  $\mathbf{u}$  は、活性化関数  $a(\cdot)$  を経て畳み込み層の出力  $\mathbf{y} = (y_{ijl})$  となる。出力  $\mathbf{y}$  は式 (3.5) で算出される。これにより、1 つのフィルタ  $\mathbf{w}$  につき、入力  $\mathbf{x}$  と横方向、縦方向、奥行きサイズが同じ  $S_x \times S_y \times S_z$  の 1 チャンネルの出力  $\mathbf{y}$  を得る。また、同様のフィルタを  $N'$  個用意し、それぞれ独立に式 (3.5) の計算を行えば、 $N'$  チャンネル分の  $S_x \times S_y \times S_z$  の出力、すなわち  $S_x \times S_y \times S_z \times N'$  のサイズの  $y_{ijl}$  を得る。これは、次の層への入力となる。



### 3.1 3D Convolutional Neural Network (3D-CNN)

$$y_{ijl} = a(u_{ijl}) . \quad (3.5)$$

本研究では、活性化関数として Rectified Linear Unit (ReLU) 関数を用いる。ReLU 関数は、入力が 0 を超えていれば入力値をそのまま出力し、0 以下ならば 0 を出力する関数である [8]。ReLU 関数は、式 (3.6) で表される。

$$a(x) = \begin{cases} x & (x > 0) \\ 0 & (x \leq 0) . \end{cases} \quad (3.6)$$

この関数は出力計算や勾配計算が高速にできるという利点がある。また、 $\max(0, x)$  という関数の性質上、多くの値が 0 となることで出力が疎になり、勾配も 0 になりやすい。これにより、深いニューラルネットワークであっても誤差は消失せずに伝播するため、一般的な誤差逆伝播法で勾配消失問題に対処できる。

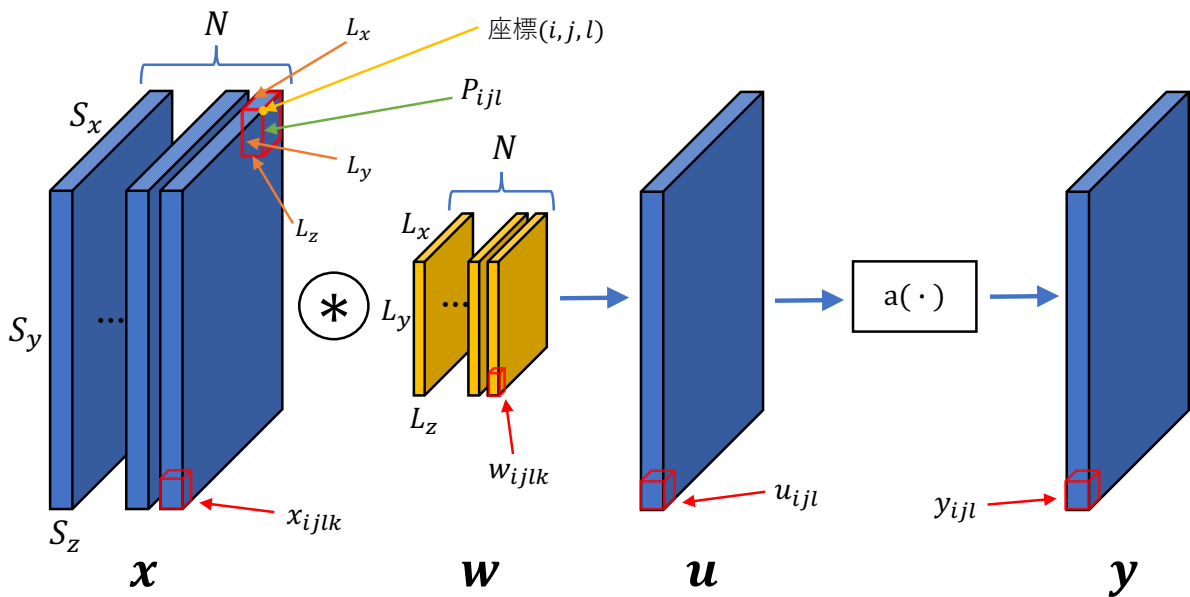


図 3.3: 畳み込み層の基本構造

## 3.1 3D Convolutional Neural Network (3D-CNN)

### 3.1.2 プーリング層

プーリング層 (Pooling Layer) は畳み込み層と対で使われ、基本的には畳み込み層の出力がプーリング層への入力となる。そのため、入力は  $S_x \times S_y \times S_z \times N$  の形をとる。

プーリング層のノード  $(i, j, l)$  は、畳み込み層と同様に、その入力側の層に局所受容野  $P_{ijl}$  を持つ。このノード  $(i, j, l)$  の出力は、 $P_{ijl}$  の内部のノード  $(p, q, r) \in P_{ijl}$  の出力  $y_{pqr}$  を集約したものである。なお、 $P_{ijl}$  のサイズは通常、畳み込み層のフィルタサイズとは無関係に設定される。また入力が複数チャンネルある場合、チャンネルごとに以上の処理を独立して行うのが一般的である。そのため、畳み込み層の出力チャンネル数とプーリング層の出力チャンネル数は一致する。

基本的にプーリングの処理は画像の縦横  $(i, j, l)$  方向に間引いて行うため、2 以上のストライド  $s$  を設定する。 $s = 2$  にしたとき、出力は入力の縦横半分のサイズになる。 $s$  と  $P_{ijl}$  のサイズによっては、隣接する出力ノードの局所受容野  $P_{ijl}$  が互いに重なり合うこともある。こうして、プーリング層の出力ノード数は入力ノード数の  $(1/s)^2$  倍になる。

プーリングの主な方法としては、平均プーリング (Average Pooling) と最大プーリング (Max Pooling) がある。本研究では、 $P_{ijl}$  に属するノードからの入力  $x_{pqrk}$  の最大値を出力  $y_{ijlk}$  とする最大プーリングを用いる。最大プーリングは式 (3.7) で表される。図 3.4 はストライドを 2 にした場合の最大プーリングの処理を表したものである。

$$y_{ijlk} = \max_{(p,q,r) \in P_{ijl}} x_{pqrk} . \quad (3.7)$$

### 3.2 Batch Normalization

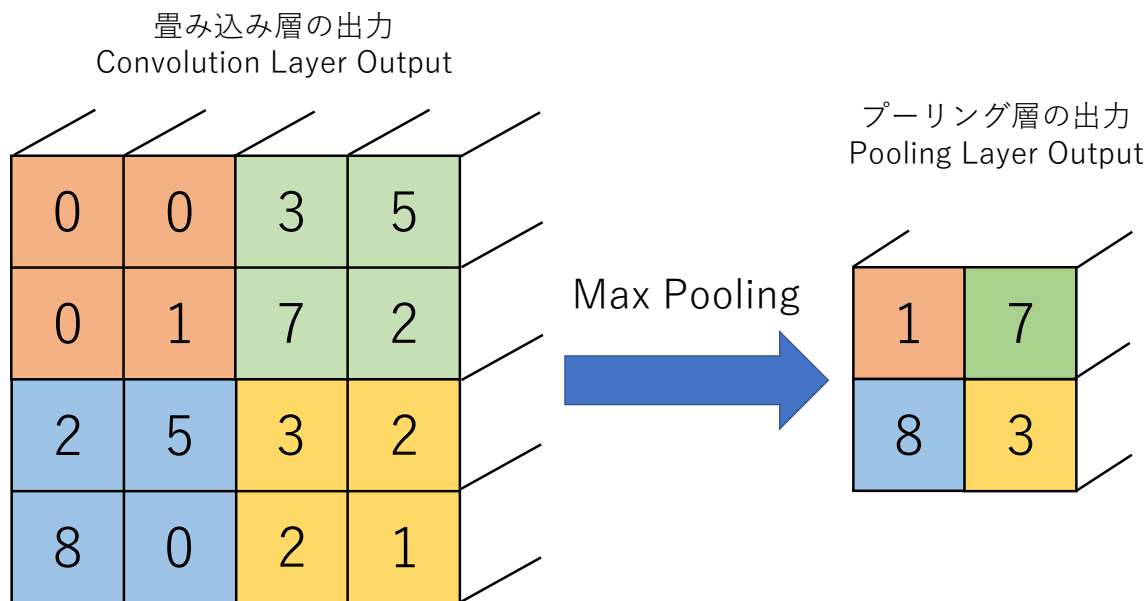


図 3.4: プーリング層

## 3.2 Batch Normalization

深いニューラルネットワークの学習が遅くなる理由として、学習が進むにつれ各層の入力が変化していく問題がある。この問題は内部共変量シフト (Internal Covariate Shift) が起きていることが挙げられる。内部共変量シフトは、学習時とテスト時の入力の分布が異なる場合に分布の変化が特徴として捉えられ、間違った方向に学習が進んでしまうことである [7]。この問題を抑制する手法として Batch Normalization が提案されており、多くの研究者や技術者に広く使われている。

Batch Normalization は、学習を行う際のミニバッチを単位として、ミニバッチごとにデータの分布の平均が 0 で分散が 1 になるように正規化を行う [9]。Batch Normalization のアルゴリズムとしては式 (3.8) より、まずミニバッチとして  $B$  という  $m$  個の入力データの集合  $\{x_1, x_2, \dots, x_m\}$  に対して、 $B$  の平均  $\mu_B$ 、 $B$  の分散  $\sigma_B^2$  を求める。その後、入力データを平均が 0 で分散が 1 のデータ  $\hat{x}_i$  になるように正規化する。このとき、 $B$  のサイズが小さいときに零除算が発生することを防ぐため、小さい実数  $\epsilon$  ( $10^{-7}$  程度) を足している。

### 3.3 Dropout

これにより、適切な入力分布にすることができる。

$$\begin{aligned}\mu_B &= \frac{1}{m} \sum_{i=1}^m x_i, \\ \sigma_B^2 &= \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2, \\ \hat{x}_i &= \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}.\end{aligned}\tag{3.8}$$

さらに式 (3.9) より、Batch Normalization は正規化されたデータに対して、固有のスケールとシフトで変換を行う。ここで、 $\gamma$  と  $\beta$  をパラメータとし、 $\gamma = 1$ ,  $\beta = 0$  からスタートした時、学習によって適した値に調整される。これによって、内部共変量シフトを抑制している。

$$y_i = \gamma \hat{x}_i + \beta.\tag{3.9}$$

### 3.3 Dropout

ニューラルネットワークはモデルの表現力が高いため、過学習しやすい問題がある。そこで様々な過学習制御方法が提案されているが、その中でも Dropout は過学習を防ぐ強力な学習方法である [10]。Dropout は複数のモデルを学習し、これらの予測結果の幾何平均を利用することで過学習を防ぐことを目的としている。しかし、複数のモデルを利用するのは学習、予測ともに計算コストが大きくなる問題がある。この問題に対し Dropout は図 3.5 のように、毎回ランダムに一定割合  $0 \leq \alpha < 1$  のノードを消した状態で学習を行う。この  $\alpha$  は層ごとに設定されるハイパーパラメータである [7]。

Dropout の例としては、ある層にノードが  $\{1, 2, 3, \dots\}$  とあったとき、最初のトレーニングでは  $\{1, 3, 5, 7, 9, \dots\}$ 、2 回目では  $\{1, 2, 4, 6, 7, \dots\}$  といったように半数程度のノードを消した部分的なニューラルネットワークで学習を行う。また、消したノードではそのトレーニングに対してパラメータの更新を行わない。学習が終了したら、各モデルの予測結果の幾

### 3.3 Dropout

何平均を最終的な予測結果とする。しかし、学習時すべての学習結果を保存し、それらの幾何平均を計算することは計算コストが高い。そのため、Dropout では予測時は各ノードからの出力を  $\alpha$  倍にした上で予測を行うことで計算コストの問題を解決した。これにより、各学習時のモデルからの予測をした上で幾何平均をとった場合の解が得られ、過学習を防ぐことができる。

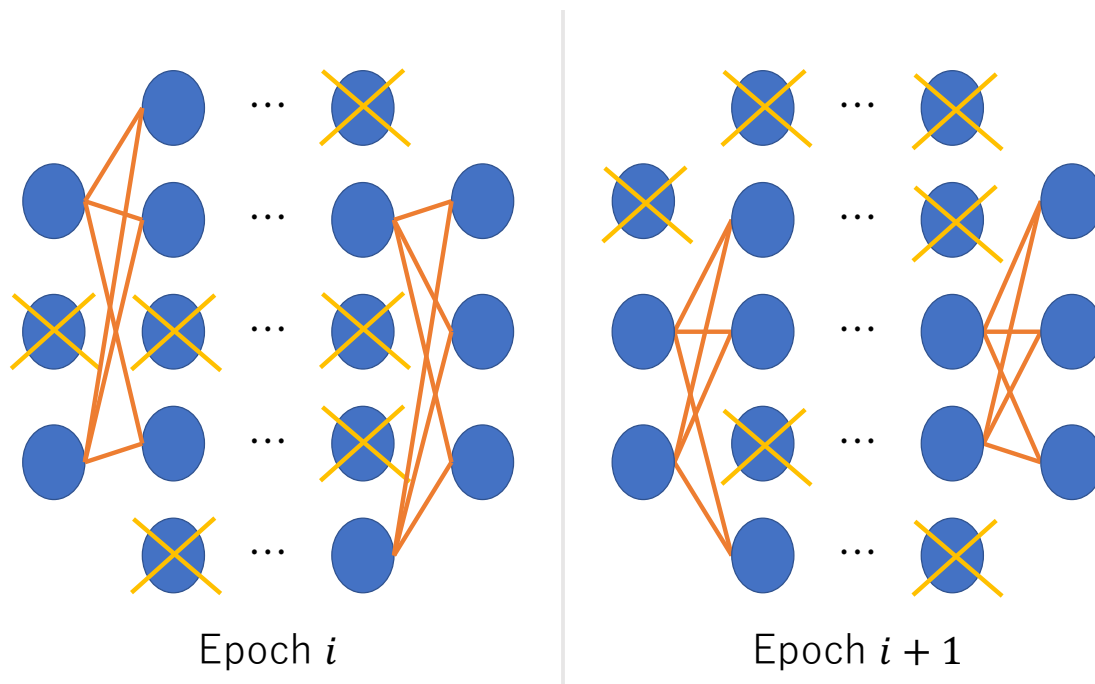


図 3.5: Dropout

## 第 4 章

# 提案手法と実験内容

本研究では，笹谷らの研究 [2] の年齢推定タスクに対して，年齢の割合が少ないデータを複製によって補正することで推定精度の向上が可能であるか検証する．

本章では，提案手法であるデータ補正手法，実験で用いたデータセット，実験で用いた年齢推定モデル，年齢推定モデルの評価手法について述べる．

### 4.1 データ補正手法

Wang らの研究 [3] では，トレーニングデータのうち年齢の割合が少ないデータを画像変換によって補正することで高い精度の年齢推定を行っている．このことから，トレーニングデータにおける年齢分布の偏りを少なくすることが年齢推定をする上で効果的であると考えられる．そこで本研究では補正手法として，トレーニングデータを複製することで年齢分布の偏りを少なくする方法を提案する．データセットでは 20 代と 80 代のデータが少なく，推定モデルが学習できない可能性があるため 20 代と 80 代のデータを除去する．また，年齢層の偏りを少なくするために年齢階級を表 4.1 の階級にする．年齢階級はスタージェスの公式 [11] を参考にし，階級数を  $k$ ，データ総数を  $N$  として式 (4.1) より決定した．その後，トレーニングデータの中で最も度数が多い階級を基準とし，他の階級のデータを基準の階級と同じ度数になるように複製する．

$$k = \log_2 N + 1 . \quad (4.1)$$

## 4.2 データセット

表 4.1: 年齢の度数分布表

年齢	30-35	36-41	42-46	47-52	53-57	58-63	64-68	69-74	75-79	Total
度数	8	22	42	55	65	67	50	35	11	355

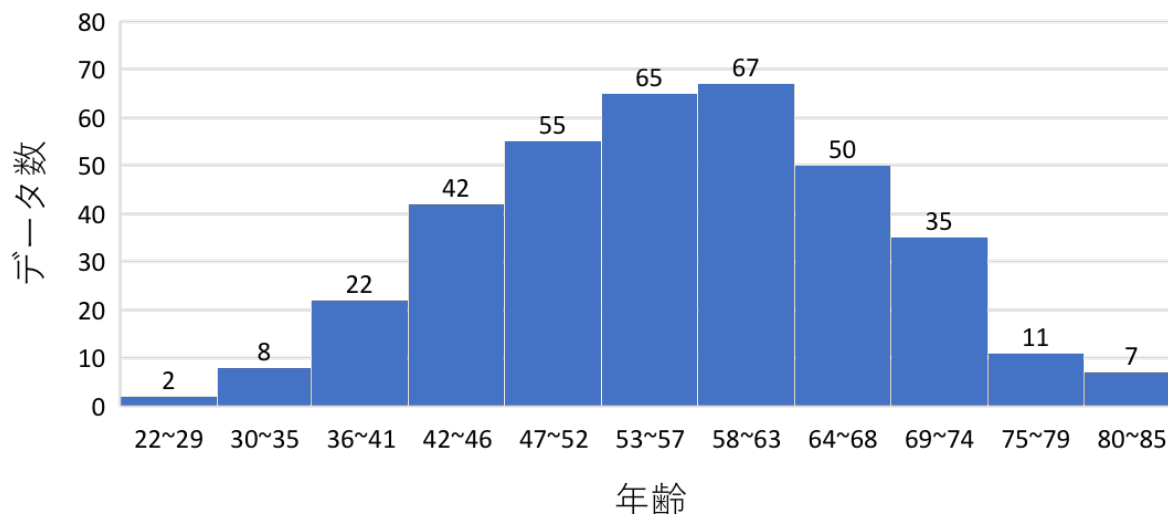


図 4.1: データセットの年齢分布

## 4.2 データセット

本研究では、実際の健康な脳ドック検診者の MRI 灰白質画像をデータとして用いた。検診者の人数は男性 207 人、女性 157 人の計 364 人である。年齢範囲は 22 歳から 85 歳であり、平均年齢 56.55 歳、標準偏差 11.11 となっている。また、20 代と 80 代のデータを除去した場合のデータは 355 人となり、平均年齢 56.23 歳、標準偏差 10.40 となっている。データの年齢分布は図 4.1 のようになっている。

使用する画像は T1 構造画像であり、前処理として SPM12 を用い、脳画像を標準脳に合わせる標準化と脳画像から脳領域分割を行うセグメンテーションを行っている。

### 4.3 推定モデル

推定モデルは入力画像が3次元構造のため、笹谷らの研究 [2] で提案された3D-CNNモデルを参考にしたものを使用する。使用した推定モデルでは、データの過学習を防止するためにBatch NormalizationとDropoutを加えている。損失関数には平均絶対誤差 (Mean Absolute Error)、活性化関数にはReLU関数、最適化手法にはAdamを用い、epoch数は100とする。

推定モデルの概要は図4.2のようになっている。また、3D-CNNのネットワークの詳細は表4.2であり、畳み込み層をConv (出力空間の次元数, 畳み込みのサイズ), Batch NormalizationをBN, ReLU関数をReLU, プーリング層をMaxPooling (プーリングのサイズ), 全結合層をFullConnection (出力空間の次元数), DropoutをDropout (ノード消す割合) として表記している。

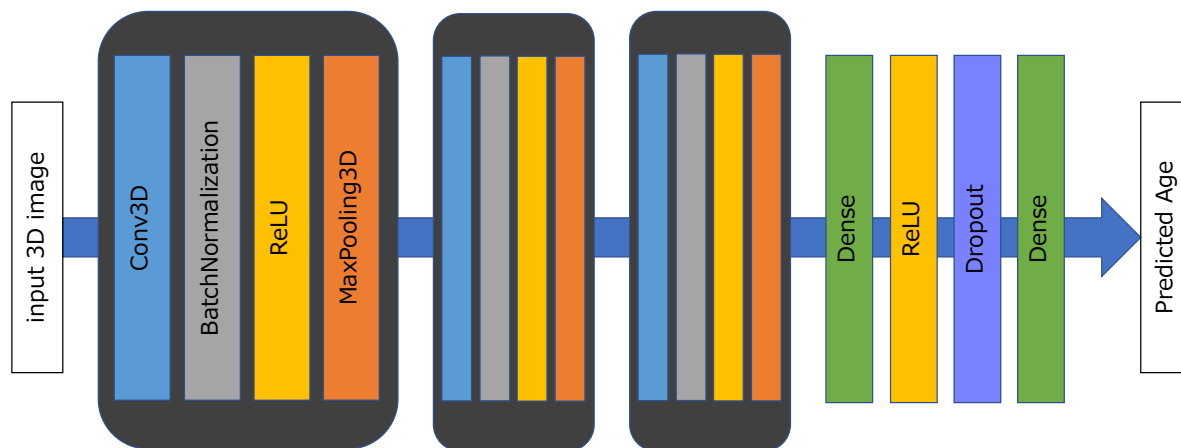


図 4.2: 年齢推定モデル



### 4.3 推定モデル

表 4.2: ネットワーク構成

Block No.	Architecture
1	Conv1(20, 4) → BN → ReLU → MaxPooling(8)
2	Conv2(20, 3) → BN → ReLU → MaxPooling(4)
3	Conv3(20, 2) → BN → ReLU → MaxPooling(2)
4	FullConnection(3000) → ReLU → Dropout(0.3) → FullConnection(1)

## 4.4 評価手法

### 4.4 評価手法

推定モデルの評価関数は平均絶対誤差 (Mean Absolute Error) を用いて行う。平均絶対誤差はデータ数を  $N$ 、予測値を  $\hat{y}_i$ 、正解値を  $y_i$  としたとき、式 (4.2) で求めることができる。

$$\text{MAE} = \frac{\sum_{i=1}^N |\hat{y}_i - y_i|}{N}. \quad (4.2)$$

また、最終的なモデルの評価としては、データのばらつきによる評価値の変動を避けるために K-分割交差検証法 (K-fold cross-validation) を用いる。K-分割交差検証法では、図 4.3 のように全データセットをランダムに  $K$  個のデータセットに分割する。その後、 $K - 1$  個をトレーニングデータとして使用し、残りの 1 個をテストデータとして使用する。これを  $K$  回繰り返すことで  $K$  個の学習済みモデルと評価値を得ることができ、得られた評価値  $E_i$  の平均  $E$  を式 (4.3) より算出し、モデルの性能とすることにより、データのばらつきによる評価値の変動を避けることができる [12]。

$$E = \frac{\sum_{i=1}^K E_i}{K}. \quad (4.3)$$

本研究では 5 分割の交差検証法を用いて 5 回モデルを学習し、算出された評価値の平均値により評価する。

#### 4.4 評価手法

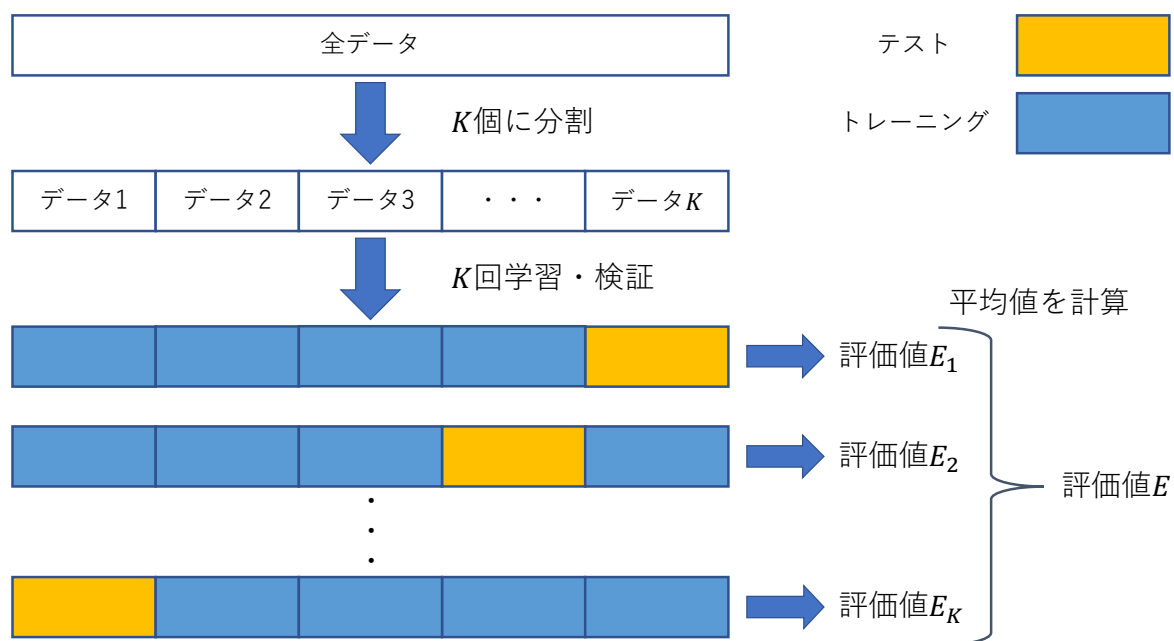


図 4.3: K-分割交差検証法

## 第 5 章

# 結果

年齢推定精度は表 5.1 より，補正なしで MAE は 9.63 歳（標準偏差 6.81），提案手法で MAE は 8.96 歳（標準偏差 6.58）となり，補正なしと比較して提案手法の MAE は 0.67 歳低くなった．表では平均値を Ave.，標準偏差を S.D. と表記している．

年齢推定の正解値と推定値の散布図 5.1 より，補正なしでは笹谷らの結果 [2] と同様にデータの平均年齢 56.55 歳付近で推定するモデルとなった．提案手法ではデータの平均年齢と関係なく幅広い年齢層で推定するモデルとなった．相関は補正なしでは 0.21，提案手法では 0.29 となり，提案手法の方が相関が高くなった．

表 5.1: 年齢推定精度

	1 回目	2 回目	3 回目	4 回目	5 回目	Ave.	SD
補正なし	11.73	8.17	8.91	9.28	10.06	9.63	6.81
提案手法	7.37	13.30	9.17	8.22	6.76	8.96	6.58

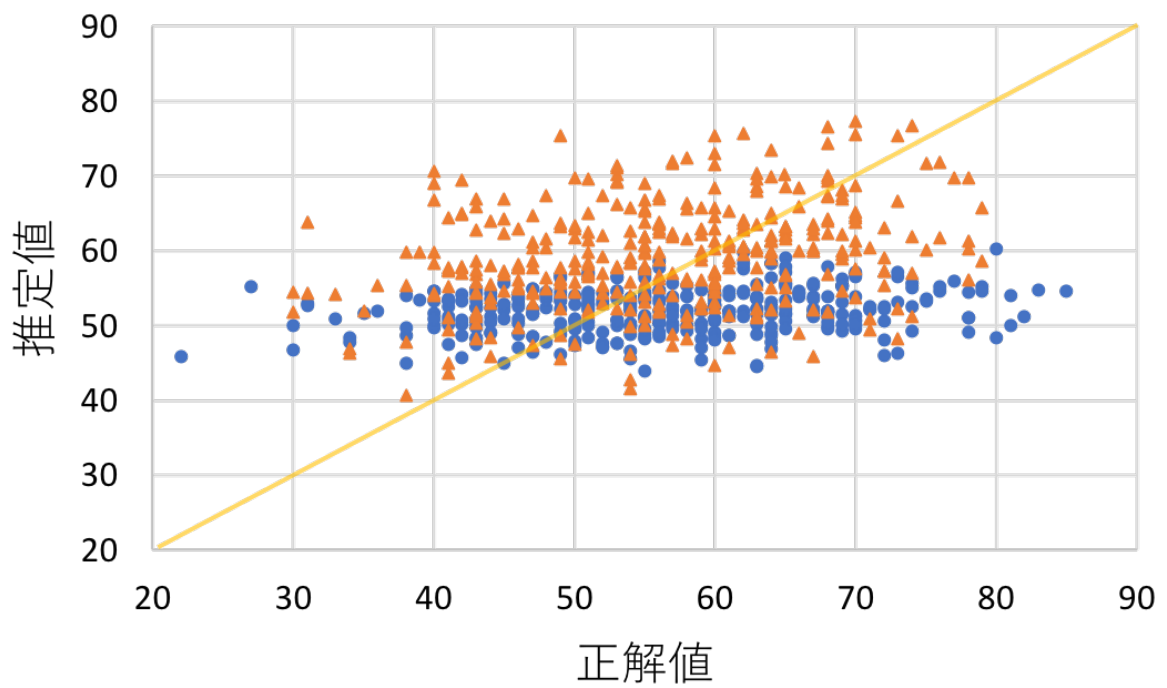


図 5.1: 補正なしと提案手法の散布図

○ : 補正なし (相関 : 0.21)

△ : 提案手法 (相関 : 0.29)

## 第 6 章

# 考察

結果より、提案手法ではデータの平均年齢を推定してしまう問題が解消されたことが分かる。しかし、補正なしの結果と比較すると全年齢層において高年齢層で推定していた。この結果の原因として、データ補正により多く複製した 70 代のデータを過学習してしまい、他の年齢層の推定に影響を与えたと考えられる。しかし、70 代と同様に多く複製した 30 代のデータを過学習していないことから、若年層の灰白質画像から年齢に関する情報を得ることは難しいのではないかと考える。

訓練データの複製によるデータ分布補正で年齢推定精度の向上が見られなかったことや、Perez らの研究 [13] からデータ分布補正の手法として画像変換によるデータ拡張が有用であることから、年齢推定においても画像変換によるデータ分布補正が有用であると考えられる。そのため、今後はどのような画像変換が年齢推定において有用であるか検証する必要があると考える。

## 第7章

### 結論

本研究では、笹谷らの研究 [2] の年齢推定タスクに対して、トレーニングデータの年齢分布をデータの複製によって補正することで年齢推定精度の向上が可能であるか検証した。結果としては、データの年齢の平均値付近で推定値を出力していた補正なしと比較して、トレーニングデータを複製によって補正した場合には、幅広い年齢層で推定値を出力するようになった。しかし、全体的に高年齢層の推定値が多くなってしまったため、データ補正により多く複製した70代のデータを過学習してしまい、他の年齢層の推定に影響を与えたと考えられる。

本研究で提案した複製によるデータ分布補正では年齢推定精度の向上が見られなかったため、今後の課題としては画像変換によるデータ分布補正を検証することが挙げられる。また、どのような画像変換が年齢推定において有用であるかも検証していく必要がある。

# 謝辞

本研究を進めるに当たり、ご指導して頂きました高知工科大学情報学群 吉田真一准教授に心から感謝致します。研究に関して右も左も分からない状態だった私でしたが、研究活動において大切なことをご教授して頂くとともに研究内容について多くのアドバイスを頂き、サポートして下さったおかげで無事に論文執筆を終えることができました。研究室活動の ITnews では、機械学習分野だけではなくネットワークなどの他分野に関しても非常に興味深いお話を聞くことができ、大変勉強になりました。また、私がネットワーク分野に興味があることから、情報学群のネットワークに関わらせて頂いたことや研究室での NAT 設置をさせて頂いたことは、大学生活において貴重な経験となりました。吉田研究室で過ごした2年間で多くのことを学び、成長することができたと感じています。社会に出た後も吉田研究室で学んだこと忘れず、これからの人生に活かしていきたいと思えます。

また、高知工科大学情報学群 四宮友貴助教には、多くのアドバイスを頂くとともに研究に関する様々な相談を聞いてくださり、無事に研究活動を進めることができました。深く感謝致します。

そして、本研究の副査を引き受けて頂いた高知工科大学情報学群 篠森敬三教授ならびに高知工科大学情報学群 中原潔教授に深く感謝致します。卒業研究発表会の質疑応答では、全く異なる視点での考え方を知ることができ、研究に役立てることができました。

同研究室の皆様にも大変お世話になりました。同期である学士4年の方々には、2年間お世話になりました。皆様とは研究に関する話だけでなく、様々な内容の話をすることができ、自分とは異なる考え方を聞くことができたため今後の人生で参考にさせて頂きたいと思えます。今春からそれぞれの選択した道に進みますが、皆様お互いに次のステージでも頑張っていけたらなと思えます。2年間ありがとうございました。また、修士の方々や学士3年の方々にも研究室活動や普段の研究室生活を含め、大変お世話になりました。

最後に、大学で4年間学ぶ機会を与えてくださり、大学生活を最後まで支えて下さった



## 謝辞

家族，共に大学生活を送った友人，研究に関わってくださった皆様に深く感謝致します。皆様のおかげで最後まで楽しく大学生活を送ることができました。この4年間私を支えてくださった皆様へ心からの感謝の気持ちと御礼を申し上げたく，謝辞に変えさせていただきます。

## 参考文献

- [1] T. Kaufmann, D. van der Meer, N. T. Doan, et al, “Common brain disorders are associated with heritable patterns of apparent aging of the brain,” *Nature Neuroscience*, Vol.22, pp.1617–1623, 2019.
- [2] 笹谷健文, “機械学習による脳構造 MRI からの個人属性の推定,” 平成 30 年度高知工科大学修士学位論文, 2019.
- [3] J. Wang, M. J. Knol, A. Tiulpin, F. Dubost, M. de. Bruijne, M. W. Vernooij, H. H. H. Adams, M. A. Ikram, W. J. Niessen and G. V. Roshchupkin, “Gray Matter Age Prediction as a Biomarker for Risk of Dementia,” *PNAS*, Vol.116, No.42, pp.21213–21218, 2019.
- [4] K. Fukushima, S. Miyake, T. Ito, “Neocognitron: A neural network model for a mechanism of visual pattern recognition,” *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.SMC-13, No.5, pp.826–834, 1983.
- [5] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, “Backpropagation Applied to Handwritten Zip Code Recognition,” *Neural Computation*, Vol.1, No.4, pp.541–551, 1989.
- [6] S. Ji, W. Xu, M. Yang, K. Yu, “3D Convolutional Neural Networks for Human Action Recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.35, No.1, pp.221–231, 2013.
- [7] 神嶋敏弘, 麻生英樹, 安田宗樹, 前田新一, 岡野原大輔, 岡谷貴之, 久保陽太郎, ポレガラダヌシカ, “深層学習 —Deep Learning—,” 近代科学社, pp.156–160, 136, 139–140, 2015.
- [8] V. Nair, G. E. Hinton, “Rectified Linear Units Improve Restricted Boltzmann Machines,” *International Conference on Machine Learning*, pp.807–814, 2010.

## 参考文献

- [9] S. Ioffe, C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *Proceedings of Machine Learning Research*, Vol.37, pp.448–456, 2015.
- [10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research*, Vol.15, pp.1929–1958, 2014.
- [11] H. A. Sturges, “The Choice of a Class Interval,” *Journal of the American Statistical Association*, Vol.21, No.153, 1926, pp.65–66, 1926.
- [12] S. Raschka, V. Mirjalili, “Python Machine Learning Second Edition,” Packt Publishing, pp.191–192, 2017.
- [13] L. Perez, J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” In: Stanford University research report, 2017.