明治大学総合数理学部

2023 年度

卒　業　研　究

合成アルゴリズムの秘匿属性推定攻撃に対する安全性評価

学位請求者　先端メディアサイエンス学科

谷口輝海

目次

第 1 章

はじめに

近年，データの利活用の活発化に伴い，データに含まれる個人のプライバシー保護が重要な課題となっている. データのプライバシーを保護する枠組みとして匿名化 [10] や差分プライバシー [9] といった技術が提案されている. しかし，これらの技術では，十分なプライバシーを提供するために，しばしばデータの有用性が著しく損なわれることがある. そこで，合成データ技術が注目されている.

合成データとは，実在するデータと同じ統計的性質を有するアルゴリズムで生成された架空のデータである. 合成データはプライバシーを保護しながら実データに近い有用性を保つことができる手段として期待されているが，プライバシーの保護度合いは明確に定量化されておらず，どのような合成アルゴリズムにおいて有用性や安全性が高く保証されるかは定かではない.

そこで，本研究では Python のオープンソースライブラリとして Synthetic Data Vault[1] で提供されている，3 つの合成アルゴリズム (Conditional Tabular GAN[2], Tabular Variational Auto Encoder[2], CopulaGAN[3]) について，分布誤差，データ列間の相関の誤差，回帰モデルにおける信頼区間の重複度合いによる有用性評価指標で，生成される合成データの有用性を定量化する. 次に，安全性について，(1) 合成データからの属性推論成功率を考慮したリスク評価指標と，(2)mehnaz らによって提案された，機械学習モデルの入出力の情報から学習データの属性値を推論する攻撃である Confidence Score based Model Inversion Attack[7] を，用いた開示リスクを評価する. CSMIA を，オリジナルのデータで学習したモデルと，合成データで学習したモデルのそれぞれに適用し，攻撃精度の差を明らかにすることを目的とする.

本研究のシステム構成図を図 1.1 に示す. このシステム構成図では，CSMIA を用いて合成データを評価するための手順を示している. まず，オリジナルデータセット D を合成アルゴリズムに入力し，合成データセット D' を得る. 次に，D と D' を用いて分類ニューラルネットワーク f_{orig} と f_{synth} を学習させる. f_{orig} と f_{synth} を用いてオリジナルデータの変数 x_2 を推定する. 続いて，f_{orig} と f_{synth} に，オリジナルデータのレコードを x_2 の可能な値で繰り返したデータを入力する. 最後に，それぞれの出力に対して CSMIA を適用することで，x'_2 を推定する.

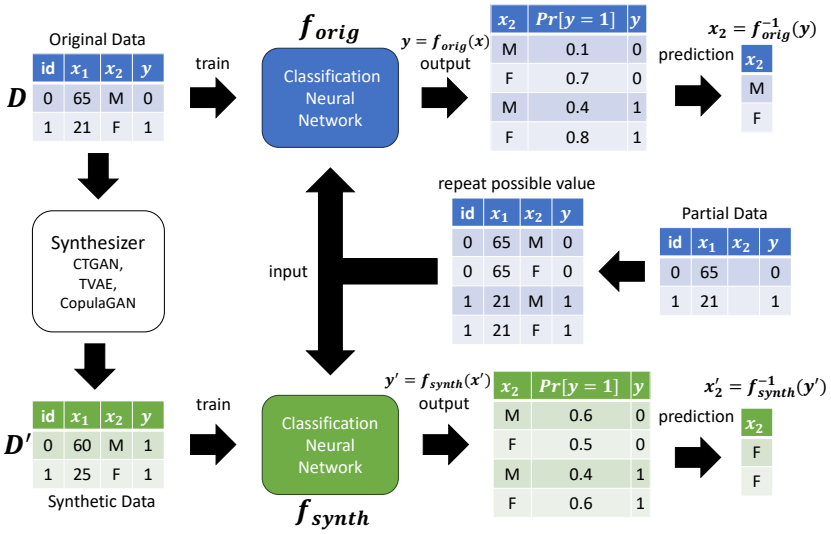
Original Data
D
id x_1 x_2 y
0 65 M 0
1 21 F 1
train
f_{orig}
Classification Neural Network
$y = f_{orig}(x)$
output
x_2 $Pr[y=1]$ y
M 0.1 0
F 0.7 0
M 0.4 1
F 0.8 1
$x_2 = f_{orig}^{-1}(y)$
prediction
x_2
M
F
Synthesizer
CTGAN,
TVAE,
CopulaGAN
input
repeat possible value
id x_1 x_2 y
0 65 M 0
0 65 F 0
1 21 M 1
1 21 F 1
Partial Data
id x_1 x_2 y
0 65 0
1 21 1
D'
id x_1 x_2 y
0 60 M 1
1 25 F 1
Synthetic Data
train
Classification Neural Network
f_{synth}
$y' = f_{synth}(x')$
output
x_2 $Pr[y=1]$ y
M 0.6 0
F 0.5 0
M 0.4 1
F 0.6 1
$x_2' = f_{synth}^{-1}(y')$
prediction
x_2
F
F

図 1.1　システム構成図

第 2 章

先行研究

2.1 Claire Little らの研究

多くの統計期間において, 無作為に抽出した国勢調査のサンプルデータを, Statistics Disclosure Control(SDC) に基づき，10% 未満のサンプリング割合で公開している.

Claire ら [11] は，公開されているサンプルデータから，異なる割合でサンプリングデータを作成し，合成データと比較することで，合成データの有用性と開示リスクを評価する枠組みを提案している. また，それによって，合成データと同様の有用性と開示リスクを持つサンプリング割合を調査した. 合成データの作成には，決定木ベースで合成データを生成するアルゴリズムを実装した R パッケージである Snthpop[17]，Xu らによって提案された CTGAN[2]，ベイジアンネットワークを用いて合成データを作成し，差分プライバシーによってプライバシーを保証する DataSynthesizer[16] の 3 つを用いた.

4 つの国勢調査データを用いて実験を行なった結果，サンプルデータの有用性とリスクの関係は曲線的であり，合成データと比較した場合の結果はデータによって異なることが示された. また，差分プライバシーを保証すると，プライバシー予算が大きくても，リスクよりも先に有用性が大幅に減少することがわかった.

2.2 Khan らの研究

khan ら [12] は，合成器と合成データ生成時のパラメータの選択が有用性に与える影響と，有用性評価指標と機械学習を用いた解析精度の相関関係を調査した.

実験には Synthpop で提供されるいくつかの合成器を使用し，有用性評価指標として，Kulback-Leibler Divergence，Confidence Interval Overlap[4] を使用した.

その結果，合成器の選択，生成データセット数，変数の合成順序が合成データの有用性に影響を与えることを示した。また，解析精度との相関について，Kulback-Leibler Divergence は相関が弱く，CIO では相関が強いということが実験的に示された.

第 3 章

準備

3.1 Synthetic Data Vault[1]

本研究では，datacebo 社が開発したオープンソースである Synthetic Data Vault ライブラリ [1] で提供される 3 つの合成アルゴリズム，CTGAN，TVAE，CopulaGAN を対象とする．

3.1.1 CTGAN[2]

CTGAN[2] は，Generative Adversarial Network[13] の一種であり，表形式のデータを生成するように設計されている．通常の GAN と同様に，生成器 (Generator) と識別器 (Discriminator) と呼ばれる 2 つのニューラルネットワークを競合させ，同時に学習させることで，学習データに類似したデータを生成するモデルを構築する．

表形式データ生成における大きな課題点として，連続値属性の分布の複雑性と離散値であるカテゴリ出現頻度の不均衡の 2 つが挙げられる．CTGAN では，Mode-Specific Normalization と Training-by-Sampling という 2 つの手法を用いて表形式データ生成特有の問題点に対処している．Mode-Specific Normalization では，任意の複雑な分布を持つ連続値属性を，複数のガウス分布の線形和として表現し，分布のモードごとに数値を正規化する．Training-by-Sampling では，各カテゴリの対数頻度に応じて，識別器に入力するデータと，生成器に入力する条件ベクトルをサンプリングすることで，全ての取り得る離散値を均等に学習させることができる．

3.1.2 TVAE[2]

Xu らによって提案された TVAE[2] は，表形式のデータを生成するために VAE を派生させたモデルである．CTGAN と同様の前処理を行い，通常の VAE と同じく学習を行う．

3.1.3 CopulaGAN[1, 3]

CopulaGAN[1, 3] は CTGAN と Copula[3] を組合わせた表形式データ生成モデルである．Copula は，多変量の累積分布関数と周辺分布関数の関係を表す関数であり，確率変数間の多様な依存関係を表現する．

3.2 相関値

表形式データでは，量的データと質的データが混在しているため，本研究では，データ形式に応じて，相関係数 (量的)，相関比 (量的と質的の組)，クラメルの連関係数 (質的) をそれぞれ相関の指標として用いる．

3.2.1 相関比 [6]

相関比 [6] は，量的変数と質的変数の関係の強さを表す指標である．0 から 1 の値を取り，1 に近いほど変数間の関連が強い．

C を取り得るカテゴリの集合とし，各カテゴリの変数の個数を n_i 個 ($i \in C$) とする．また，$x_{i,j}$ をカテゴリ i であるもののうち，j 番目の量的変数値とし，量的変数の平均値を $\overline{x}$，質的変数の各カテゴリ毎の i 番目の量的変数の平均値を $\overline{x_{i,j}}$ とする．このとき，相関比 η^2 は以下の式で計算される．

$$\eta^2 = \frac{\sum_{i \in C} n_i (\overline{x}_i - \overline{x})^2}{\sum_{i \in C} \sum_{j=1}^{n_i} (x_{ij} - \overline{x})^2}$$

表 3.2 の例における **age**(量的変数) と **marriage**(質的変数) を用いて，

$$\overline{x_{age}} = 40.8, \overline{x}_{age,Single} = 34, \overline{x}_{age,Marital} = 47, \overline{x}_{age,Divorced} = 42$$

のとき，

$$\begin{aligned}\eta^2 &= \frac{2 \cdot (\overline{x}_{Single} - \overline{x})^2 + 2 \cdot (\overline{x}_{Marital} - \overline{x})^2 + 1 \cdot (\overline{x}_{Divorced} - \overline{x})^2}{(43 - \overline{x})^2 + (29 - \overline{x})^2 + (25 - \overline{x})^2 + (42 - \overline{x})^2 + (65 - \overline{x})^2} \\ &\fallingdotseq 0.066\end{aligned}$$

である．

3.2.2 クラメルの連関係数 [8]

クラメルの連関係数は質的変数同士の関連の強さを表す指標である．0 から 1 の値を取り，1 に近いほど関連が強いとされる．質的変数のカテゴリの数をそれぞれ r, c としたとき，r 行 c 列のクロス集計表を作成し，その χ^2 値から算出される．サンプルレコード数を n とすると，連関係数 V は，

$$V = \sqrt{\frac{\chi^2}{n \cdot \min(r-1, c-1)}}$$

で定義される．表 3.2 の例における **marriage**(質的変数) と **sex**(質的変数) を用いた計算例を示す．サンプルデータにおけるクロス集計表は表 3.1 のような 3×2 の表となる．ここから χ^2 値を算出すると，$\chi^2 \fallingdotseq 3.3$ となり，

$$V = \sqrt{\frac{3.3}{5 \cdot \min(3-1, 2-1)}} = 0.66$$

表 3.1　D のクロス集計

marriage \ sex	**Male**	**Female**	total
Single	2	1	2
Marital	0	2	2
Divorced	1	0	1
total	3	2	5

表 3.2　サンプルデータ D

id	**age**	**height**	**marriage**	**sex**	**income**
0	43	170	Single	Female	$> 50K$
1	29	161	Marital	Female	$\leq 50K$
2	25	179	Single	Male	$\leq 50K$
3	42	174	Divorced	Male	$> 50K$
4	65	153	Marital	Female	$> 50K$

表 3.3　合成サンプルデータ D'

id	**age**	**height**	**marriage**	**sex**	**income**
0	40	175	Single	Male	$\leq 50K$
1	35	155	Single	Male	$> 50K$
2	25	171	Marital	Female	$\leq 50K$
3	42	165	Divorced	Female	$> 50K$
4	52	145	Marital	Female	$\leq 50K$

3.3　Confidence Interval Overlap[4]

合成データの有用性を測るために，Karr らによって提案されている Confidence Interval Overlap (CIO) を使用する. オリジナルデータ D と合成データ D' でそれぞれ構築した回帰モデルにおいて，回帰係数の信頼区間に占める重複区間の割合の平均で計算される. u_o, ℓ_o と u_s, ℓ_s をそれぞれ，オリジナルデータ D と合成データ D' に対する回帰係数の信頼区間の上界と下界とするとき，CIO は，

$$CIO = \frac{1}{2}\left(\frac{\min(u_o, u_s) - \max(l_o, l_s)}{u_o - l_o} + \frac{\min(u_o, u_s) - \max(l_o, l_s)}{u_s - l_s}\right)$$

と定める. 信頼区間が完全に一致する場合，CIO は最大値 1 をとる. 本研究では重複がない場合の CIO 値を 0 とする.

例えば，あるデータについて，図 3.1 に示すようにオリジナルデータが $[1, 5]$，合成データが $[3, 7]$ となる信

頼区間が与えられたとする. このとき，CIO の値は，

$$
\begin{aligned}
CIO &= \frac{1}{2}\left(\frac{\min(5,7)-\max(1,3)}{5-1}+\frac{\min(5,7)-\max(1,3)}{7-3}\right)\\
&= 0.5
\end{aligned}
$$

である.

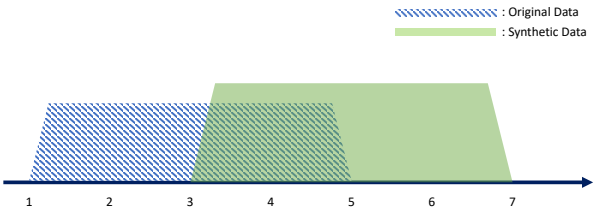

図 3.1　信頼区間の例

3.4　Targeted Correct Attribution Probability[5]

Taub らによって提案された Targeted Correct Attribution Probability(TCAP) は，合成データを公開したときに元データの値がどれだけ漏洩するかのリスクを評価するための指標である. TCAP の算出において，敵対者は合成データにアクセスできると仮定し，質的変数をいくつか得る. 1 つを target 変数，残りを target 変数を予測するための key 変数とする. オリジナルデータと合成データにおける key 変数と target 変数をそれぞれ K_o, T_o, K_s, T_s とする. このとき，TCAP を計算するために，合成データの各レコード j について，レコード j がもつ K_s の組合わせと同じレコードのうち，target 変数がレコード j と一致するものの割合を算出する. これは，Within Equivalece Class Attribution Probability (WEAP) と呼ばれ，

$$
\begin{aligned}
WEAP_{s,j} &= \Pr(T_{s,j}|K_{s,j})\\
&= \frac{Pr[T=T_{s,j}, K=K_{s,j}]}{Pr[K=K_{s,j}]}\\
&= |\{(K_{s,i}, T_{s,i}) \mid T_{s,i}=T_{s,j}, K_{s,i}=K_{s,j}, i\in[n]\}|
\end{aligned}
$$

と定義される.

$WEAP$ の値を閾値として，TCAP を計算するレコードを定める. 本研究では WEAP の値を 1 とした. つまり，$WEAP_{s,j}=1$ であるような合成レコード j について TCAP の値を算出する. 対応するオリジナルデータのレコード j について，TCAP を.

$$
\begin{aligned}
TCAP_{o,j} &= \Pr(T_{s,j}|K_{s,j})_o\\
&= \frac{\sum_{i=1}^{n}[T_{o,i}=T_{s,j}, K_{o,i}=K_{s,j}]}{\sum_{i=1}^{n}[K_{o,i}=K_{s,j}]}\\
&= \frac{Pr[T=T_{o,j}, K=K_{s,j}]}{Pr[K=K_{s,j}]}
\end{aligned}
$$

とする.

$WEAP_{s,j} = 1$ に対応するオリジナルレコードが存在しない場合，分母は 0 となるため TCAP は定義されない. TCAP は 1 に近いほど開示リスクが高く，0 に近いほど開示リスクは低い.

表 3.2 のデータセットを用いて表 3.3 のような合成データセットが得られたとする． $K_s = \{\textbf{marriage}, \textbf{sex}\}, T_s = \{\textbf{income}\}$ としたとき，合成データにおいて key 変数の組合せは $id = 0, 1$ が (*Single*, *Male*), $id = 3$ が (*Divorced*, *Female*), $id = 2, 4$ が (*Marital*, *Female*) となり，3 パターン存在する. ここで，$id = 0, 1$ については target 変数の値がそれぞれ異なるため $WEAP_{s,()} = 0.5$ である. $id = 2, 3, 4$ については target 変数の値が一意に定まっているため，$WEAP = 1$ となる. $WEAP = 1$ となるレコードの key 変数に一致するレコードをオリジナルデータ D(表 3.2) から抜き出すと，$id = 1, 4$ が該当する. このうち，target 変数が一致するレコードは $id = 1$ であるので，TCAP の値は 0.5 となる.

3.5 Confidence Score based Model Inversion Attack[7]

Mehnaz らによって提案された Confidence Score based Model Inversion Attack(CSMIA) は，学習済み分類モデルの出力値から学習データの属性推論を行うアルゴリズムである. 攻撃者は，標的モデルに入力を行うことができ，分類結果と各クラスの所属確率にアクセスできるものとする. 分類モデルの学習に用いたデータを $x_1, x_2, \ldots, x_d, y$ とする. ただし，x_i は説明変数，y は目的変数 (クラスラベル) である. このとき，攻撃者は $x_2, \ldots, x_d, y$ に関する情報を有しており，あるセンシティブな属性 x_1 について推論を行う.

CSMIA では，モデルに対し正しい属性値で入力を行なった場合に，矛盾のない，より高い確信度で正しい予測結果を返すという考えに基づき，属性を推論する. まず，攻撃者はセンシティブ属性について可能な値を全て列挙した入力用のデータを作成し，モデルに対して入力を実行する. 例えば，x_1 が **Marital** であるとき，考えられる値は Married と Single の 2 つであり，攻撃者は機械学習モデルに対して，$(Married, x_2, \ldots, x_d)$ と $(Single, x_2, \ldots, x_d)$ を入力し，出力 $\hat{y}_{Married}$ と $\hat{y}_{Single}$ を得る. そして，以下の 3 ケースについて，センシティブ属性値を推論する.

(1) $\hat{y}_{Married} = y$ かつ $\hat{y}_{Single} \neq y$，もしくは，$\hat{y}_{Married} \neq y$ かつ $\hat{y}_{Single} = y$ のとき，y について正しい出力をした入力を推論値とする. 例えば，$\hat{y}_{Married} = y$ かつ $\hat{y}_{Single} \neq y$ であれば $\hat{x_1} = Married$ を返す.

(2) $\hat{y}_{Married} = y$ かつ $\hat{y}_{Single} = y$ のとき
y の所属確率が高い方を推論値とする.

(3) $\hat{y}_{Married} \neq y$ かつ $\hat{y}_{Single} \neq y$ のとき
y の所属確率が低い方を推論値とする.

第 4 章

実験方法

4.1　実験目的

SDV ライブラリで提供される 3 つの合成アルゴリズムの有用性と開示リスクを比較評価するために，様々な評価指標を用いて実験を行う.

4.2　使用データ

本研究では，Adult データセット [14] を用いた. 表 4.1 に Adult データセットの概要を示す.

表 4.1　Adult データセットの概要

変数名	データタイプ	カテゴリ数	詳細
age	連続	-	年齢
workclass	質的	7	職業クラス
fnlwgt	連続	-	重み（Final Weight の略）
education	質的	16	教育レベル
marital.status	質的	7	結婚の状態
occupation	質的	14	職業
relationship	質的	6	家族関係
race	質的	5	人種
sex	質的	2	性別
capital.gain	連続	-	資産利益
capital.loss	連続	-	資産損失
hours.per.week	連続	-	週の労働時間
native.country	質的	41	出身国
income	質的	2	収入（$\leq 50K, > 50K$）

4.3 有用性評価

4.3.1 分布評価

データセットの属性ごとに頻度分布を求め，オリジナルデータと合成データとの間で各カテゴリ値の出現確率の誤差を求めた. ただし，数値変数に関しては，階級を 16 個に分割して計算した. 誤差の算出には Mean Abusolute Error(MAE) を用いる.

4.3.2 相関評価

データセット内の各属性の組について相関値を計算し，オリジナルデータと合成データとの間で MAE を求めた. なお，データセット内には量的変数と質的変数が混在しているため，組合せる属性のデータ形式に応じて相関指標を選択している. 量的変数間に相関係数, 量的変数と質的変数に相関比, 質的変数間にクラメルの連関係数を用いた.

4.3.3 CIO

オリジナリデータと合成データの類似度を CIO を用いて評価する. ロジスティック回帰を使用し，各回帰係数に対する CIO の基本統計量を算出する. 説明変数には age，workclass，education，hours.per.week，race を使用し，目的変数は income とする.

4.4 リスク評価

4.4.1 TCAP

各合成アルゴリズムで生成した合成データについて TCAP の値を算出し，開示リスクを評価する.

表 4.2 に各合成データで TCAP 値算出のために用いた変数を示す. key 変数の数を 3 ~ 5 個とし，key 変数と target 変数の全組合せ 96 通りについて TCAP 値を算出した.

表 4.2　使用変数とカテゴリの数

変数	workclass	relationship	race	marital.status	sex	income
カテゴリ数	7	6	5	7	2	2

4.4.2 CSMIA

オリジナルデータで学習した機械学習モデルと，オリジナルデータを用いて生成した合成データで学習した機械学習モデルに対し，CSMIA を実行し，機械学習モデルの精度と攻撃の精度の関係を調査する.

攻撃対象のモデルには，income の 2 値分類を行うニューラルネットワーク (NN) をオープンソースライブラリ PyTorch[15] で実装した. NN は 3 層の全結合層で構成され，中間層の活性化関数には ReLU 関数，出力層の活性化関数には sigmoid 関数を用いた. 最適化アルゴリズムは Adam を採用し，学習率は 0.005 としている. また，推論対象の変数を marital.status とし，7 つのカテゴリ値の内，Married-civ-spouse, Married-spouse-absent,

Married-AF-spouse を Married，Divorced, Never-married, Separated, Widowed を Single としてまとめて 2 値変数とした.

第 5 章

結果と考察

5.1　実験結果

5.1.1　有用性評価

頻度分布

図 5.1 に各属性のカテゴリ値出現確率の平均絶対誤差 (MAE) を示す. ここで，CTGAN の値でソートしている.

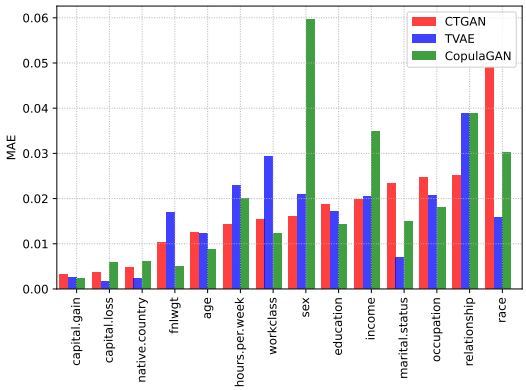

図 5.1　分布誤差

相関評価

図 5.2 に各合成器の相関指標ごとの相対絶対誤差 (MAE) を示す. ここで，CTGAN の値でソートしている.

CIO

表 5.2 に各合成データで算出した回帰係数の上限と下限の値を示す. 表 5.2 における Overlap は，1 つの変数から算出した回帰係数における CIO の値である. 表 5.1 に各合成データとオリジナルデータとの間で計算した CIO の結果を示す. なお，CIO を計算するために使用したロジスティック回帰の F1-score を表中に示している.

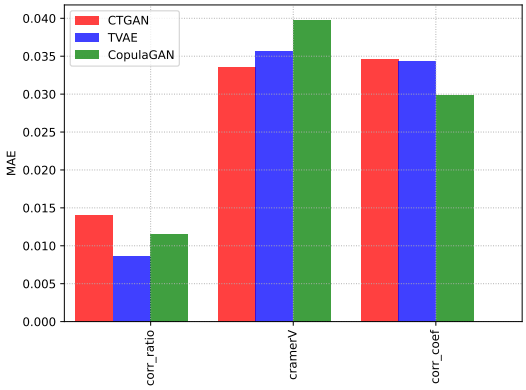

図 5.2　平均相関誤差

表 5.1　CIO 値の基本統計量

合成器	F1	最小値	中央値	最大値	平均値	標準偏差
オリジナル	0.66	-	-	-	-	-
CTGAN	0.64	0.00	0.62	0.92	0.56	0.31
TVAE	0.65	0.00	0.00	0.91	**0.16**	**0.27**
CopulaGAN	0.69	0.00	0.41	0.88	0.40	0.36

表 5.2　各合成器における回帰係数毎の CIO

	orig_lower	orig_upper	CTGAN_lower	CTGAN_upper	TVAE_lower	TVAE_upper	CopulaGAN_lower	CopulaGAN_upper	CTGAN_Overlap	TVAE_Overlap	CopulaGAN_Overlap
Intercept	5.856	6.809	4.696	5.607	2.495	3.730	4.166	4.929	0.000	0.000	0.000
workclass(T.Local-gov)	0.447	0.777	0.339	0.751	2.387	3.501	0.998	1.337	0.829	0.000	0.000
workclass(T.Private)	0.380	0.655	0.293	0.594	2.077	3.172	0.848	1.116	0.747	0.000	0.000
workclass(T.Self-emp-inc)	−0.305	0.061	−0.348	0.087	0.382	1.593	0.221	0.589	0.921	0.000	0.000
workclass(T.Self-emp-not-inc)	0.713	1.039	0.454	0.796	2.614	3.725	0.940	1.244	0.249	0.000	0.316
workclass(T.State-gov)	0.607	0.978	0.546	0.935	2.532	3.751	1.395	1.743	0.864	0.000	0.000
workclass(T.Without-pay)	−0.320	3.938	2.171	3.996	**−155720.310**	**155765.088**	0.527	2.255	0.692	0.500	0.703
education(T.11th)	−0.556	0.175	0.320	1.390	1.485	2.570	0.146	0.815	0.000	0.000	0.042
education(T.12th)	−1.026	−0.130	−0.764	0.361	−2.448	0.746	−0.665	0.036	0.636	0.640	0.680
education(T.1st-4th)	−0.010	1.714	−0.471	0.606	0.158	1.585	0.098	1.156	0.465	0.914	0.807
education(T.5th-6th)	−0.173	0.932	−0.695	0.368	0.382	1.109	0.055	0.984	0.499	0.627	0.869
education(T.7th-8th)	−0.040	0.790	−0.340	0.442	1.005	1.647	−0.020	0.491	0.598	0.000	0.807
education(T.9th)	−0.123	0.843	0.047	1.011	0.802	2.293	−0.031	0.540	0.825	0.035	0.796
education(T.Assoc-acdm)	−2.107	−1.506	−1.653	−0.964	–	–	−1.755	−1.272	0.228	–	0.464
education(T.Assoc-voc)	−2.028	−1.441	−2.254	−1.573	**−62471.400**	**62515.005**	−1.783	−1.368	0.721	0.500	0.704
education(T.Bachelors)	−2.713	−2.165	−2.680	−2.048	−0.982	−0.628	−2.231	−1.840	0.877	0.000	0.145
education(T.Doctorate)	−3.777	−3.076	−3.451	−2.673	−2.096	−1.618	−3.562	−3.068	0.509	0.000	0.839
education(T.HS-grad)	−1.369	−0.823	−1.321	−0.688	0.259	0.615	−0.767	−0.374	0.849	0.000	0.000
education(T.Masters)	−3.069	−2.500	−3.229	−2.565	−1.998	−1.567	−2.240	−1.812	0.823	0.000	0.000
education(T.Preschool)	−0.531	3.497	−0.907	3.115	**−24382.939**	**24426.347**	−0.920	0.842	0.906	0.500	0.560
education(T.Prof-school)	−3.886	−3.226	−4.189	−3.464	−2.541	−1.959	−3.808	−3.303	0.611	0.000	0.883
education(T.Some-college)	−1.756	−1.206	−1.321	−0.665	0.082	0.442	−1.108	−0.716	0.192	0.000	0.000
race(T.Asian-Pac-Islander)	−0.791	−0.045	−0.717	−0.150	–	–	−1.158	−0.564	0.880	–	0.343
race(T.Black)	−0.277	0.434	−0.012	0.554	0.359	0.754	−0.139	0.469	0.707	0.149	0.874
race(T.Other)	−0.647	0.390	0.000	0.767	0.158	2.539	−0.202	0.520	0.443	0.160	0.696
race(T.White)	−0.896	−0.215	−1.103	−0.572	−0.296	−0.018	−1.235	−0.670	0.543	0.206	0.365
age	−0.049	−0.044	−0.020	−0.015	−0.041	−0.036	−0.027	−0.022	0.000	0.000	0.000
hours_per_week	−0.043	−0.038	−0.036	−0.031	−0.066	−0.059	−0.041	−0.036	0.000	0.000	0.581

5.1.2 TCAP

key 変数を 3 ~ 5 個としたときの全組合せについて，TCAP 値を算出した. 表 5.3 に各合成器ごとの TCAP 値の基本統計量を示す.

表 5.3　各合成データにおける TCAP 値の基本統計量

合成器	key 変数の数	組合せ数	最小値	最大値	平均値	標準偏差
CTGAN	3	60	0.000	1.000	**0.602**	0.363
	4	30	0.301	0.977	**0.762**	0.152
	5	6	0.694	0.939	0.798	0.108
TVAE	3	60	0.340	0.995	**0.810**	0.149
	4	30	**0.682**	0.970	**0.855**	0.089
	5	6	**0.800**	0.958	**0.890**	0.071
CopulaGAN	3	60	0.000	1.000	0.666	0.350
	4	30	0.462	0.970	0.779	0.136
	5	6	0.622	0.926	**0.786**	0.118

keys	target	TCAP_CTGAN	TCAP_TVAE	TCAP_CopulaGAN
relationship + race + marital.status	workclass	0.800	0.835	0.697
workclass + race + marital.status	relationship	0.471	0.578	0.750
workclass + relationship + marital.status	race	0.579	0.783	0.667
workclass + relationship + race	marital.status	0.901	0.928	0.925
relationship + race + sex	workclass	0.714	0.865	0.905
workclass + race + sex	relationship	0.000	0.622	1.000
workclass + relationship + sex	race	0.250	0.832	0.636
workclass + relationship + race	sex	0.822	0.964	0.952
relationship + race + income	workclass	0.625	0.640	0.600
workclass + race + income	relationship	0.889	0.422	1.000
workclass + relationship + income	race	0.697	0.826	0.889
workclass + relationship + race	income	0.960	0.937	0.923
relationship + marital.status + sex	workclass	0.500	0.853	0.727
workclass + marital.status + sex	relationship	0.000	0.743	0.000
workclass + relationship + sex	marital.status	0.724	0.938	0.846
workclass + relationship + marital.status	sex	0.558	0.956	0.658
relationship + marital.status + income	workclass	0.700	0.856	0.643
workclass + marital.status + income	relationship	0.375	0.782	0.550
workclass + relationship + income	marital.status	0.522	0.980	0.667
workclass + relationship + marital.status	income	0.926	0.966	0.930
relationship + sex + income	workclass	0.000	0.656	0.000
workclass + sex + income	relationship	1.000	0.696	0.500
workclass + relationship + income	sex	0.983	0.965	1.000
workclass + relationship + sex	income	0.974	0.934	0.900
race + marital.status + sex	workclass	0.875	0.802	0.897
workclass + marital.status + sex	race	1.000	0.797	0.735
workclass + race + sex	marital.status	0.471	0.634	1.000
workclass + race + marital.status	sex	0.625	0.856	0.684
race + marital.status + income	workclass	0.667	0.543	0.917
workclass + marital.status + income	race	0.833	0.881	0.962
workclass + race + income	marital.status	0.727	0.795	0.824

workclass + race + marital.status	income	0.908	0.928	0.927
race + sex + income	workclass	0.000	0.802	0.000
workclass + sex + income	race	0.000	0.875	1.000
workclass + race + income	sex	0.828	0.882	0.840
workclass + race + sex	income	0.804	0.762	0.750
marital.status + sex + income	workclass	0.000	0.340	0.000
workclass + sex + income	marital.status	1.000	0.829	1.000
workclass + marital.status + income	sex	0.746	0.907	0.643
workclass + marital.status + sex	income	0.895	0.885	0.902
race + marital.status + sex	relationship	0.500	0.491	0.000
relationship + marital.status + sex	race	0.000	0.730	0.909
relationship + race + sex	marital.status	0.000	0.995	0.000
relationship + race + marital.status	sex	0.550	0.815	0.700
race + marital.status + income	relationship	0.800	0.821	0.429
relationship + marital.status + income	race	0.800	0.893	0.857
relationship + race + income	marital.status	0.940	0.926	0.941
relationship + race + marital.status	income	0.983	0.971	0.963
race + sex + income	relationship	0.000	0.756	0.000
relationship + sex + income	race	0.000	0.896	0.000
relationship + race + income	sex	0.500	0.867	0.500
relationship + race + sex	income	0.968	0.967	0.975
marital.status + sex + income	relationship	0.000	0.813	0.000
relationship + sex + income	marital.status	0.000	0.883	0.000
relationship + marital.status + income	sex	1.000	0.701	0.846
relationship + marital.status + sex	income	0.949	0.974	0.971
marital.status + sex + income	race	1.000	0.935	0.909
race + sex + income	marital.status	0.000	0.614	0.000
race + marital.status + income	sex	0.800	0.521	0.524
race + marital.status + sex	income	0.982	0.962	0.966
relationship + race + marital.status + sex	workclass	0.728	0.825	0.769
workclass + race + marital.status + sex	relationship	0.301	0.772	0.609
workclass + relationship + marital.status + sex	race	0.692	0.784	0.694
workclass + relationship + race + sex	marital.status	0.891	0.905	0.843
workclass + relationship + race + marital.status	sex	0.685	0.938	0.823
relationship + race + marital.status + income	workclass	0.717	0.792	0.681
workclass + race + marital.status + income	relationship	0.707	0.779	0.657
workclass + relationship + marital.status + income	race	0.683	0.830	0.772
workclass + relationship + race + income	marital.status	0.935	0.969	0.925
workclass + relationship + race + marital.status	income	0.953	0.958	0.930
relationship + race + sex + income	workclass	0.688	0.718	0.786
workclass + race + sex + income	relationship	0.547	0.712	0.647
workclass + relationship + sex + income	race	0.800	0.854	0.791
workclass + relationship + race + income	sex	0.929	0.967	0.944
workclass + relationship + race + sex	income	0.958	0.938	0.918
relationship + marital.status + sex + income	workclass	0.667	0.833	0.615
workclass + marital.status + sex + income	relationship	0.752	0.844	0.462
workclass + relationship + sex + income	marital.status	0.627	0.970	0.722
workclass + relationship + marital.status + income	sex	0.863	0.957	0.858
workclass + relationship + marital.status + sex	income	0.931	0.966	0.921
race + marital.status + sex + income	workclass	0.750	0.682	0.846
workclass + marital.status + sex + income	race	0.819	0.835	0.816
workclass + race + sex + income	marital.status	0.600	0.726	0.868
workclass + race + marital.status + income	sex	0.729	0.862	0.664
workclass + race + marital.status + sex	income	0.908	0.926	0.916
race + marital.status + sex + income	relationship	0.692	0.848	0.500
relationship + marital.status + sex + income	race	0.842	0.781	0.778
relationship + race + sex + income	marital.status	0.923	0.920	0.970

relationship + race + marital.status + income	sex	0.581	0.791	0.679
relationship + race + marital.status + sex	income	0.977	0.969	0.965
relationship + race + marital.status + sex + income	workclass	0.694	0.800	0.684
workclass + race + marital.status + sex + income	relationship	0.710	0.863	0.622
workclass + relationship + marital.status + sex + income	race	0.713	0.820	0.761
workclass + relationship + race + sex + income	marital.status	0.910	0.957	0.866
workclass + relationship + race + marital.status + income	sex	0.820	0.943	0.859
workclass + relationship + race + marital.status + sex	income	0.939	0.958	0.926

表 5.4: 変数の組合せと TCAP 値

5.1.3 CSMIA

実験に使用した NN の分類精度を表 5.5 に示す. なお，合成データで学習した機械学習モデルのテストデータには，オリジナルデータを分割して作成したテストデータを使用している. 各データで学習した NN に対して，CSMIA を実行した結果を表 5.6 に示す. ここで，CSMIA においては，オリジナルモデルの学習データについて推論をした. 評価指標として Precision, Recall, f1-Score, Accuracy を用いた.

表 5.5　NN による目的変数の識別精度

学習データ	y (income)	Precision	Recall	f1-Score	Accuracy
オリジナル	$\leq 50K$	0.875	0.928	0.900	0.846
	$>50K$	0.731	0.597	0.657	
CTGAN	$\leq 50K$	0.851	0.932	0.890	0.826
	$>50K$	0.711	0.506	0.591	
TVAE	$\leq 50K$	0.848	0.915	0.880	0.813
	$>50K$	0.662	0.502	0.571	
CopulaGAN	$\leq 50K$	0.875	0.895	0.885	0.825
	$>50K$	0.658	0.612	0.634	

表 5.6　CSMIA による属性推定精度

合成器	sensitive Attribute	Precision	Recall	f1-Score	Accuracy
Original	Married	0.684	0.420	0.520	0.628
	Single	0.605	0.821	0.628	
CTGAN	Married	0.374	0.553	0.446	**0.340**
	Single	0.259	0.144	0.185	
TVAE	Married	0.645	0.458	0.536	0.619
	Single	0.605	0.768	0.677	
CopulaGAN	Married	0.694	0.410	0.515	0.630
	Single	0.605	0.833	0.701	

5.2 考察

5.2.1 分布/相関評価

図 5.1 で，各合成器において，最も誤差が小さくなった変数の数は CTGAN が 4 つ，TVAE が 4 つ，CopulaGAN が 6 つとなった. income や sex といった 2 値変数では CopulaGAN の誤差が顕著に大きくなっている. また, CTGAN において最も誤差が小さい変数では, カテゴリの数が少ない変数が多く, CopulaGAN において最も誤差が小さい変数では，カテゴリの数が多い変数が多い. 一方，最もカテゴリ数が多い native.country においては TVAE が誤差が最も小さく，カテゴリ出現頻度誤差についての，一貫した誤差の傾向はないように考えられる.

図 5.2 の相関比と相関係数について，CTGAN における誤差が，他の合成器に比べて大きい. 一方で，Cramer の連関係数については，CTGAN のにおける誤差が最も小さくなっており，カテゴリ変数間の相関を保持している. また，相関係数では CopulaGAN，相関比では TVAE の誤差が最も小さくなっている. このことから，相関保持の観点からは，CTGAN や CopulaGAN といった GAN をベースとした合成方式ではカテゴリ変数の合成に適しており，TVAE では数値変数の合成に適している.

ただし，元々の相関の値が小さく，誤差の値も小さいため，大きな差は見られない.

5.2.2 CIO

CTGAN において，中央値，最大値，平均値について，他の 2 つの合成器と比べて高い値を示している. 最小値に関してはいずれの合成器においても 0.00 を示しているが，表 5.2 の各合成器における Overlap の値を見ると，CIO の値が 0.00 となる変数の数が，CTGAN では 4 個，TVAE では 16 個，CopulaGAN では 9 個となり，CTGAN 以外の合成器では多くの変数において，オリジナルデータと信頼区間が全く重複しておらず，類似性が低い.

表 5.2 で，TVAE の *workclass(T.Without-pay)*，*education(T.Assoc-voc)*，*education(T.Preschool)* において，信頼区間の幅が非常に広くなっている. このとき，オリジナルデータの信頼区間を全て包含し，CIO の定義式の第一項の値が 1 となるため，結果としては CIO が 0.5 を超えるものの，オリジナルデータと全く類似していないという問題が発生する. この問題に対処するためには，信頼区間の幅の比を用いて重み付けすることなどが考えられる.

また，TVAE において，表 5.2 の *education(T.Assoc-acdm)*，*race(T.Asian-Pac-Islander)* といった変数で，カテゴリが存在せず，これらの変数について CIO が算出できない場合がある. このようにサンプリング時にカテゴリの多様性が失われたことも，CIO を下げる一因であると考えられる.

CIO では，最大値，中央値，平均値において最も高い値を示した CTGAN がオリジナルデータを最もよく近似できている. 一方で，TVAE は，表 5.1 の太字に示すように，平均値において著しく低い値を示しており，標準偏差も小さいことから，オリジナルデータをうまく近似できていない. また，CIO では，信頼区間が完全に包含される場合に必ず 0.5 以上の値を示してしまうことを考慮に入れて評価することが重要である.

5.2.3 TCAP

CTGAN と CopulaGAN では最小値，最大値，で顕著な差は見られないが，平均値では，key 変数の数が 3, 4 つのときには，CopulaGAN の TCAP 値が CTGAN よりも高く，key 変数の数が 5 つになると，CTGAN の値が CopulaGAN の値を上回る. また，表 5.3 の TVAE の行の太字に示すように，TVAE では平均値と最小値において，他の 2 つの合成器に対して，高い値を示した. TVAE における TCAP の最小値を与えた変数の組は，key 変数が *race*，*marital.status*，*sex*，*income* 及び，*relationship*，*race*，*marital.status*，*sex*，*income* で，target 変数はいずれも *workclass* である.

平均的に見ると，TVAE が最もリスクが高く，key 変数の数が少ない時には CopulaGAN のリスクが高くなると考えられる.

5.2.4 CSMIA

NN の精度では，CTGAN が正解率において他の合成器と比べて高い精度を示し，CTGAN と CopulaGAN の一部指標においてオリジナリデータで学習した NN よりも高い結果となった.

しかし，CSMIA の精度においては，CTGAN で著しく推論精度が低下しており，CopulaGAN ではオリジナルよりも高い正解率で推論に成功している.

第 6 章

結論

本研究では，SDV ライブラリで提供される 3 つの表形式データ合成アルゴリズムである CTGAN，TVAE，CopulaGAN について，データ分布や相関の誤差，CIO や TCAP といった評価指標を用いて合成データの有用性とリスクを定量的に評価した. また，学習済み機械学習モデルの入出力から学習データの属性推論を行う CSMIA を用い，合成データで学習した機械学習モデルにおいて，どれだけ属性推論リスクがあるのかを実験により明らかにした.

その結果，相関指標では，CTGAN がカテゴリカル変数間の相関保持に適しており，CopulaGAN では数値変数間の相関保持に適していることが分かった. また，CIO を用いた評価では，TVAE において著しく低い値を示し，オリジナルデータとの類似度が低いことが分かった. TCAP においては，key 変数が 3, 4 つのとき，CTGAN で低く，key 変数が 5 つのとき CopulaGAN で低い値となった. さらに，TVAE では他の二つの合成器と比べて，TCAP の平均値が key 変数の数に関わらず 0.8 以上の値を示し，リスクが高いことが分かった. CSMIA を用いた評価では，CTGAN において属性推定リスクが大幅に軽減できることを実験的に示した.

しかし，CIO の計算において，一方のデータにおける信頼区間の幅が極端に大きくなってしまった場合に，CIO 値が高く出てしまうなど，有用性の指標としての課題があると考えられ，今後の研究課題としたい.

参考文献

[1] Patki, Neha and Wedge, Roy and Veeramachaneni, Kalyan, “The Synthetic data vault”, IEEE International Conference on Data Science and Advanced Analytics (DSAA), pp.399-410, 2016.

[2] Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante, and Kalyan Veeramachaneni, “Modeling Tabular Data Using Conditional GAN”. Neural Information Processing Systems, pp.7335–7345, 2019.

[3] Yi Sun, Alfredo Cuesta-Infante, and Kalyan Veeramachaneni, “Learning vine copula models for synthetic data generation”, AAAI Conference on Artificial Intelligence, pp.5049–50572018.

[4] Karr, A.F., Kohnen, C.N., Oganian, A., Reiter, J.P., Sanil, A.P., ” A Framework for Evaluating the Utility of Data Altered to Protect Confidentiality”. The American Statistician, pp.224-232, 2012.

[5] Jennifer Taub, Mark Elliot, Maria Pampaka, Duncan Smith, “Differential Correct Attribution Probability for Synthetic Data”, An Exploration, PSD 2018, p0.122-37, 2018.

[6] Ronald A. Fisher, “STATISTICAL METHODS FOR RESEARCH WORKERS”, 1925.

[7] Shagufta Mehnaz, Sayanton V. Dibbo, Ehsanul Kabir, Ninghui Li, Elisa Bertino, “Are Your Sensitive Attributes Private? Novel Model Inversion Attribute Inference Attacks on Classification Models”, USENIX Security, pp.4579-4596, 2022.

[8] Cramér, H, “Mathematical Methods of Statistics”, Princeton University Press, pp.282. 1946.

[9] C. Dwork, “Differential privacy”, M. Bugliesi, B. Preneel, V. Sassone, and I. Wegener, (eds) Automata, Languages and Programming, volume.4052, 2006.

[10] Sweeney, L., “k-anonymity: a model for protecting privacy”, International Journal on Uncertainty, Fuzziness and Knowledge-based Systems, pp.557-570, 2002.

[11] Claire Little, Mark Elliot, Richard Allmendinger, “COMPARING THE UTILITY AND DISCLOSURE RISK OF SYNTHETIC DATA WITH SAMPLES OF MICRODATA”, PSD 2022, pp.234-249, 2022.

[12] Md Sakib Nizam Khan, Niklas Reje, Sonja Buchegger, “Utility Assessment of Synthetic Data Generation Methods”, PSD 2022.

[13] I. J. Goodfellow, et al., “Generative Adversarial Nets”, Neural Inf. Process Syst, pp2672-2680, 2014

[14] Becker Barry, and Kohavi Ronny. 1996. Adult. UCI Machine Learning Repository. https://doi.org/10.24432/C5XW20.

[15] Paszke, Adam and Gross. “Automatic differentiation in PyTorch”, NIPS-W, 2017.

[16]“DataSynthesizer: Privacy-Preserving Synthetic Datasets”, SSDBM 2017, pp.1-5, 2017.

[17] Nowok B, Raab GM, Dibben C,“synthpop: Bespoke Creation of Synthetic Data in R.”, Journal of Statistical Software, pp.1-26, 2016.

謝辞

本研究を行うにあたり，多くの方より御指導いただきました．特に，多大なる御指導を受け賜りました，明治大学総合数理学部先端メディアサイエンス学科，菊池浩明教授に深く感謝申し上げます．

大学院生の先輩の皆様には，多忙な中にも関わらず，メンターとして多くのご指導をいただきました。心より感謝申し上げます．

また，昨年度の研究においては，共同研究者である當麻君の協力とアイデアに大変助けられました．心より感謝いたします．そして，菊池研究室の同期や後輩との交流は，研究だけでなく人間関係の面でも非常に充実したものでした．ありがとうございました．

最後に，学部での学生生活を送るにあたり，ご支援いただきました全ての方々に感謝の意を表するとともに，謝辞とさせていただきます．

付録 A

歩容に基づく個人識別における Kinect と OpenPose の多人数同時個人識別精度

A.1　はじめに

人の歩き方の特徴を表す歩容は解像度の低いカメラ映像からでも取得できることから，犯罪捜査などの分野において，個人識別新たな手段として，近年注目されている．歩容に基づく属性推定・個人識別手法には，深度センサやウェアラブルデバイスなどの特定のハードウェアを本人に装着する方法 [1] や，歩容のシルエット画像列などを用いて外部から観測する方法 [2] が知られている．しかし，特定のハードウェアを用いる方法は使用場面が限られ，シルエット画像列を用いる方法では服装や髪型，携帯品などの外乱の影響を受けやすいという問題がある．加えて，街中に設置された防犯カメラ映像等を用いる際には，画角内に複数の人間が映っていることが想定されるが，既存手法では主に単独での歩行についてしか評価されていなかった．

そこで，本研究では，複数人数を同時にリアルタイムで検出する機能を持つ，Kinect[5] と OpenPose[3] に注目する．Kinect と OpenPose について，何人まで同時に識別できるか，どの程度の精度で識別できるかを明らかにすることを目的とする．本研究の個人識別の構成と流れを図 A.1 に示す．

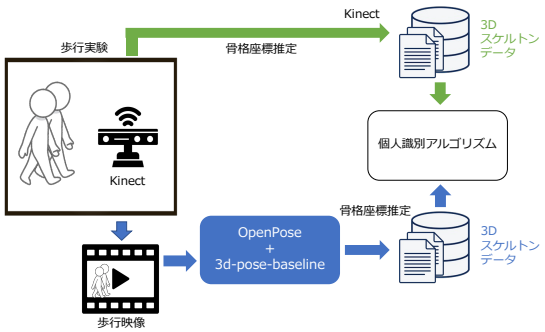

図 A.1　個人識別システム構成図

A.2　準備

A.2.1　OpenPose

OpenPose[3] は，Zhe らによって開発されたオープンソースである. 静止画像または動画からリアルタイムに複数人数の 2D 姿勢推定を行う深層学習モデルである. 姿勢推定 (Human Pose Estimation) では，人の頭部，肩，肘，手，腰，膝，足を検出し，人がどのような姿勢を取っているかを推定する. OpenPose は深度センサなどの特別な機器を必要とせず，単眼カメラのみで姿勢の推定ができ，25 点を検出できる. 図 A.2 に推定結果のプロット例を示す. 多人数を同時に検出できる利点の一方で，人数の増加に伴う推定精度の劣化は明らかではない.

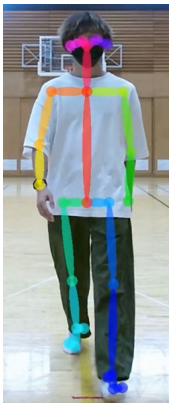

図 A.2　OpenPose の姿勢推定例

A.2.2　3d-pose-baseline

3d-pose-baseline[4] は，Julieta らによって開発されたオープンソースであり，OpenPose の出力を入力とし，深度を推定する深層学習モデルである. 2 次元画像や動画から 3 次元の姿勢を推定する.

本研究では，OpenPose と 3d-pose-baseline を用いて歩行映像から 3 次元の骨格座標データを取得する. 図 A.3 に 3d-pose-baseline の出力を 3D プロットした例を示す.

A.2.3　Kinect

Kinect[5] は Microsoft 社によって開発されたゲーム向けデバイスである．RGB カメラと深度センサ，マイクを備え，姿勢推定や音声認識を提供する．姿勢推定で検出される関節の数は一人当たり 25 点であり，手指検出や手のポーズ検出ができる．個人を追跡する機能を持ち，最大で 6 人までを同時に検知する.

本研究では，Kinect for Windows v2[5] を用いて映像と姿勢情報の取得を行った．図 A.4 に Kinect で取得した姿勢情報の 3D プロット例を示す.

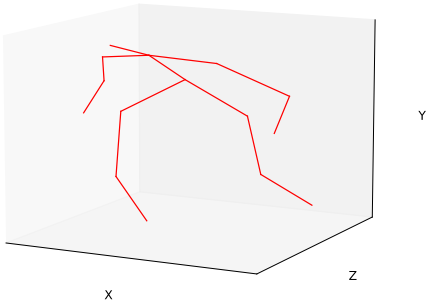

図 A.3　3d-pose-baseline の 3 次元姿勢推定例

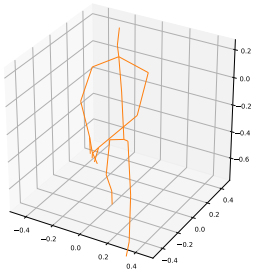

図 A.4　Kinect による 3D スケルトンデータの例

A.2.4　推定器の比較

表 A.1 に本研究で用いた姿勢推定ツール Kinect, OpenPose, 3d-pose-baseline の機能や特性を示す.

なお，本研究では OpenPose と 3d-pose-baseline を組み合わせて姿勢の 3 次元情報を取得するため，この 2 ツールをまとめて OpenPose と呼称することとする.

表 A.1　ツールの機能比較

ツール	出力	特徴	用途	原理
Kinect	カラー画像: 1920×1080 深度画像: 512 × 424 FPS:30 1 人あたり 25 関節	処理が高速なため, リアルタイムでの推論が可能. 最大で 6 人までを同時に追跡可能	リアルタイム骨格検出 (3d) 音声認識	深度センサを用いた RandomForest による推論
OpenPose	関節位置座標 2D データ 25 関節 各関節の推論信頼度	ハードウェア不要 多人数の同時検出が可能 人物追跡ができない	2D 姿勢推定	深層学習
3d-pose-baseline	関節位置座標 3D データ 16 関節	OpenPose の出力から深度を推定する	3D 姿勢推定	深層学習

A.2.5　DTW 距離

DTW（Dynamic Time Warping）距離 [7] は，2 つの時系列データ間の類似度の 1 つである．時系列データ S と T の DTW は，S の各データ点に対して，T の最小距離の点を選び，それらの距離の総和で定める．そのため，時系列の長さが異なっていたり周期がずれていたりしても類似度を与える．

図 A.5 の時系列データ $S = (0, 2, 1)$，$T = (1, 3, 0, 1)$ の例を考えよ．表 A.2 の距離行列 $dist(S, T)$ を求める．ここで，i 行 j 列（$i > 0, j > 0$）の要素 $dist[i, j]$ は

$$\begin{aligned} dist[i, j] =& cost(S[i-1], T[j-1]) \\ &+ min(dist[i-1, j], dist[i-1, j-1], dist[i, j-1]) \end{aligned}$$

とする．$cost$ 関数は 2 つのデータ点の距離を求める関数である．この例ではデータ点が 1 次元の値であるが，多次元のときは，マンハッタン距離もしくはユークリッド距離を用いて定める．このとき，求める時系列 S, T の DTW 距離は，$dist[3, 4] = 3$ である．

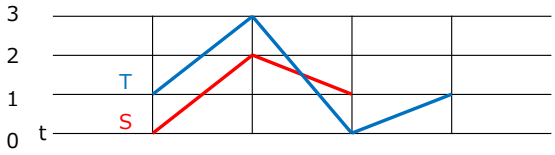

図 A.5　時系列データ S と T の例

表 A.2　距離行列 $dist(S, T)$ の例

1	∞	2	4	3	3
2	∞	2	2	4	5
0	∞	1	4	4	5
-	0	∞	∞	∞	∞
S \ T	-	1	3	0	1

A.2.6　先行研究

三好ら [8] は，Kinect を用いて取得した歩容データから特徴量を定義し，男女の平均の差から性別の識別をした．また，推定率の高い順に特徴量を統合し，99.86% の推定率を達成した．

阪田ら [9] は，歩容のシルエット画像列 GEI を入力とする CNN を構成し，年齢の推定を行なった．図 A.6 に示すように，性別や大まかな年代を推定した上で，年齢の推定を行なった．その結果，平均絶対誤差が 5.83 歳となり，既存研究を大きく上回る性能を示した．

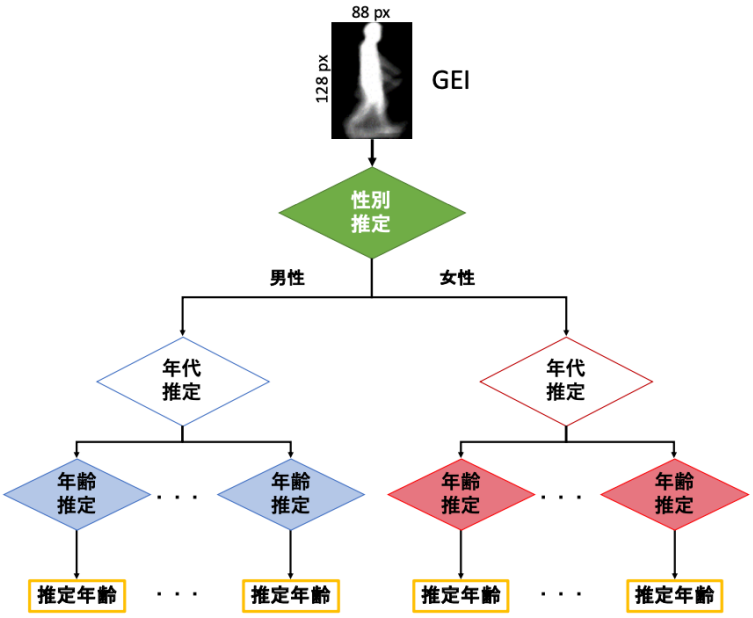

図 A.6　多段階年齢推定器のフローチャート ([9] より引用)

A.3　個人識別

A.3.1　個人識別手法

本研究では，森ら [1] の手法に従い，個人識別を行なう．森手法では，Kinect と OpenPose を用いて取得した関節の 3 次元座標をそれぞれ測定し，一歩分の時系列データの DTW 距離を算出して個人識別を行う．識別手法は以下の 4 ステップから成る．

(1) サイクル切り出し
(2) 関節座標の相対座標化
(3) DTW 距離の計算
(4) 個人識別

A.3.2　サイクル切り出し

身体の部位 ℓ の時刻 t における 3 次元空間の絶対座標を $a_\ell(t) = (x, y, z)$ とする．ここで，時刻 t の単位はフレームレートに対応する．測定時間の絶対座標の時系列データ $\langle a_\ell(t_1), a_\ell(t_2), \ldots \rangle$ から歩行の 1 サイクル分を抽出する．

まず，時刻 t の左右の足の絶対座標 $a_{LF}(t)$, $a_{RF}(t)$ から，両足の間の間隔を

$$\Delta(t) = sign \cdot \|a_{RF}(t) - a_{LF}(t)\|$$

により計算する．ここで，$sign$ は $\{-1, +1\}$ の値を取る符号であり，右足が前の状態を正とする．

次に，両足間の距離 $(\Delta(1), \ldots, \Delta(n))$ の時系列データにフーリエ変換を適用し，全周波数成分の 1/30 の低周波数成分のみを残して，残りを 0 とする．すなわち，ローパスフィルタをかけることでノイズを除去し，そのピーク間を 1 サイクルとする．

A.3.3 関節座標の相対座標化

歩行中の各関節の座標について，身体の中心付近に位置する比較的安定した関節が原点となるように相対座標化を行う．

関節 ℓ の時刻 t における絶対座標を $a_\ell(t)$，中心の関節の時刻 t における絶対座標を $a_c(t)$ とすると，相対座標 r は

$$r_\ell(t) = a_\ell(t) - a_c(t)$$

と定める．本研究において，身体の中心 c は Kinect で Spine-Mid，OpenPose で Spine(脊椎) を用いた．

A.3.4 DTW 距離の計算

各時系列データの類似度を DTW 距離を用いて定める．本研究では，2.4 節の *cost* 関数として 3 次元ベクトルのユークリッド距離

$$\|\mathbf{p}_i - \mathbf{q}_i\| = \sqrt{(p_{i,x} - q_{i,x})^2 + (p_{i,y} - q_{i,y})^2 + (p_{i,z} - q_{i,z})^2}$$

を用いる．歩行 1 サイクルの関節 ℓ の 2 つの時系列データ $R_\ell = \langle r_\ell(t_1), ..., r_\ell(t_n) \rangle$ と $R'_\ell = \left\langle r'_\ell(t_1), ..., r'_\ell(t_{n'}) \right\rangle$ の DTW 距離 $d(R, R')$ を R と R' の類似度とする．DTW 距離の性質から，$R = R'$ ならば $d(R, R') = 0$ であり，n と n' は一致する必要はない．

また，複数の関節を用いたときの類似度は次のように定める．異なる関節 m と関節 ℓ について 2 つの時系列データ (R_ℓ, R_m) と (R'_ℓ, R'_m) があるとき，統合 DTW 距離 $D((R_\ell, R_m), (R'_\ell, R'_m))$ は，ℓ と m についての DTW 距離の L2 ノルム (ユークリッド距離)，すなわち，$\sqrt{d(R_\ell, R'_\ell)^2 + d(R_m, R'_m)^2}$ とする．同様に，k 種の関節を統合した場合も，k 次元のユークリッド距離で類似度を定める．

A.3.5 個人識別

単独歩行

ある単独歩行のデータに対して，その他の単独歩行のデータ全てとの間で DTW 距離を計算し，最も DTW 距離が小さかった歩行データの該当者を識別結果とする．すなわち，サイクル切り出しと相対座標化を行った単独歩行のデータ N 組のうち i 番目の歩行データを W_i と表したとき，次の問題の解 j^* が識別結果である．

$$j^* = \underset{i \neq j \in \{1,\ldots,N\}}{\arg\min} D(W_i, W_j)$$

複数人歩行

単独歩行の識別と同様にして行う．ある複数人歩行データの識別結果は，全ての単独歩行データとの間で DTW 距離を計算し，最も DTW 距離が小さかった歩行データの該当者を表す id である．

A.4　実験

A.4.1　歩行実験

実験目的

Kinect と OpenPose の多人数同時個人識別の精度を比較するため，Kinect を用いて歩行映像と各関節座標の時系列を得る実験を行う．

実験方法

Kinect によって得た映像と関節座標を保存するためのシステムは Processing を用いて開発した．Processing で Kinect for Windows v2 を扱うためのライブラリとして KinectPV2[5] を利用した．

単独歩行と複数人歩行をそれぞれ観測した．複数人歩行は 2〜6 人が同時に歩行し，そのデータを得る実験である．実験参加者の詳細を表 A.3 に示す. 実験環境は図 A.7 の通りである．

単独歩行では Kinect に対して直進する方向とそこから ± 30° 傾いた方向の 3 パターンについて，1 人当たり 2 回ずつ測定した. 複数人歩行では，2 ~ 6 人と人数を変えながら，全員が直進するパターンを 3 回，交差が発生するパターンを 3 回測定した．

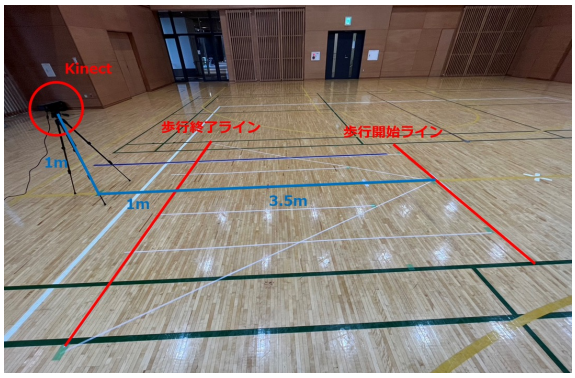

図 A.7　実験環境

表 A.3　実験参加者の情報

項目	環境
実験日	2022 年 7 月 16 日
実験時刻	9:30 から約 2 時間
場所	明治大学 中野キャンパス 多目的ホール
年齢	20 代前半
性別	男性 4 名 女性 3 名
人数	7 名

A.4.2　センシング誤り評価実験

実験目的

Kinect で推定した関節座標データに対して Kinect のセンシング誤りによるデータの誤差がどの程度存在しているのかを調べる．歩行の向きや人数の変化による Kinect の推定精度を評価する．

実験方法

データからランダムにフレームを選んで可視化し，手動でラベル付けを行う．データの選び方は単独歩行と複数人歩行で異なる．

単独歩行については，実験参加者 7 人の歩行データについて，それぞれ 3 種類の歩行方向から 3 フレームずつランダムに選ぶ．可視化するフレームの合計は $7 \times 3 \times 3 = 63$ フレーム分である．

複数人歩行については，各歩行人数のデータに対して 1 人あたり 2 フレーム，合計 $7 \times 2 = 14$ フレームをランダムに選んで可視化する．(ただし 2 人歩行については 1 人参加していない参加者が存在するため，複数人歩行の合計可視化フレーム数としては 68 フレームである．)

ラベル付けに関しては，正常に見えるもの，一部に異常が見られるもの，全体的に異常なものの 3 種類に分類する．各ラベルのサンプルを次の図 A.8, A.9, A.10 に示す．

実験結果

歩行方向を正面と正面以外に分けたときの各ラベルの占める割合を表 A.4 に示す．複数人歩行と単独歩行を合わせて歩行人数に関してラベルの割合を表 A.5 と図 A.11 に示す．

表 A.4　各歩行方向に対するセンシング状態の割合

歩行方向 (Kinect 基準)	正常	一部異常	全体異常
正面	1.0 (21 / 21)	0.0 (0 / 21)	0.0 (0 / 21)
正面以外	0.48 (20 / 42)	0.33 (14 / 42)	0.19 (8 / 42)
合計	0.65 (41 / 63)	0.22 (14 / 63)	0.13 (8 / 63)

表 A.5　歩行人数に対するセンシング状態の割合

歩行人数	正常	一部異常	全体異常
1	0.65 (41 / 63)	0.22 (14 / 63)	0.13 (8 / 63)
2	0.67 (8 / 12)	0.17 (2 / 12)	0.17 (2 / 12)
3	0.64 (9 / 14)	0.29 (4 / 14)	0.071 (1 / 14)
4	0.36 (5 / 14)	0.43 (6 / 14)	0.21 (3 / 14)
5	0.29 (4 / 14)	0.50 (7 / 14)	0.21 (3 / 14)
6	0.43 (6 / 14)	0.29 (4 / 14)	0.29 (4 / 14)

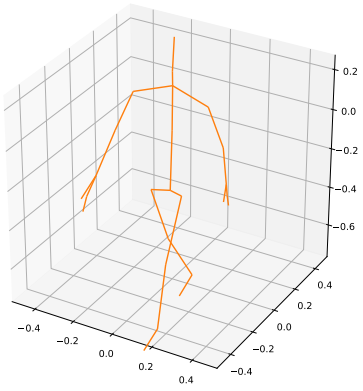

図 A.8　正常なフレーム

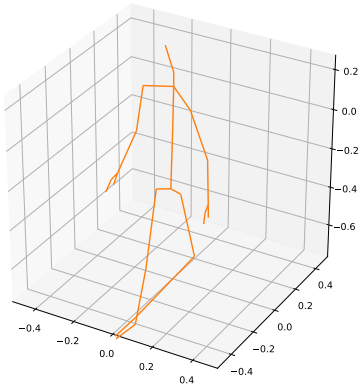

図 A.9　一部に異常が見られるフレーム

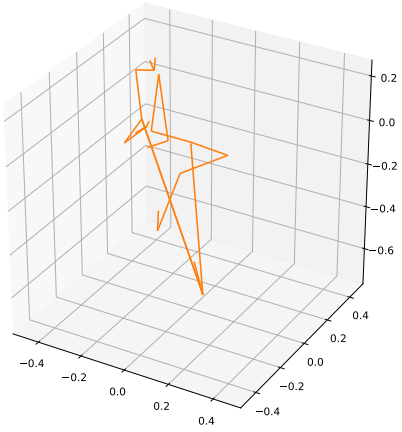

図 A.10　全体的に異常なフレーム

A.4.3　個人識別実験

実験目的

単独歩行・複数人歩行に対する Kinect と OpenPose の精度を調べる．

実験方法

3.5 節の手法に基づき，各歩行データに対して個人識別を行う．また，$1 \leq k \leq 5$ の top と topk の推定を含めた精度 top_kAcc を用いて，Kinect と OpenPose の間で精度を比較する．ただし，top_kAcc の計算においては，

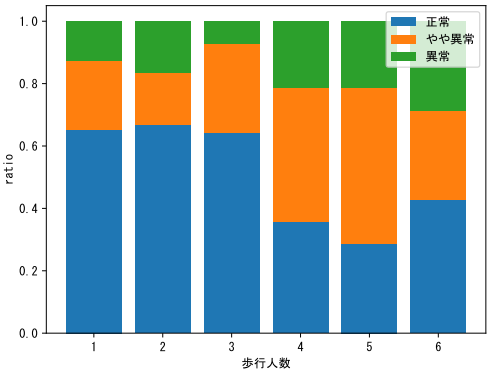

図 A.11　歩行人数変化とセンシング誤りの割合

DTW 距離が最小のものから昇順で該当個人 id を並べたものを識別結果としている．

実験結果

単独歩行における識別結果の混同行列を表 A.6, A.7 に示し，top_kAcc $(k = 1, 2, 3, 4, 5)$ を図 A.12 に示す．また，歩行人数を増加させたときの識別精度の推移を top_kAcc $(k = 1, 3, 5)$ について求め，その結果を図 A.13, A.14, A.15 に示す．

A.4.4　考察

Kinect のセンシング誤り

単独歩行について，Kinect に対する歩行角度でセンシング精度に違いが見られた．Kinect は姿勢推定を行うとき，各画素に対して 25 関節のうちどの関節に該当するかを分類し，関節ごとに分類された画素の中心座標を求めることで姿勢推定を行っている．そのため，Kinect に対して角度が付くことで，Kinect から見えづらい関節が増えて姿勢推定は失敗しやすくなると考えられる．また，歩行人数に対する Kinect のセンシング精度について，4, 5 人歩行のとき低下した．ここから，Kinect が問題なく性能を発揮できるのは 3 人同時姿勢推定までであると言える．また，6 人歩行のとき 1 3 人歩行ほどではないが 4, 5 人歩行より精度が上がった．これは，そもそも 6 人を同時に捉えられず，歩行人数が少ない状態に等しくなっているからと考えられる．実際に，6 回の 6 人歩行のうち 2 回はそもそも 6 人分の歩容が得られていなかった．

個人識別

歩行人数が 1 ~ 3 人のとき，k の値に関わらず OpenPose よりも Kinect の方が識別精度 top_kAcc が高かった．しかし，4 人歩行で精度が逆転した．6 人歩行に対しては Kinect の方がやや有利であったがあまり精度が変わらなかった．このことは，Kinect のセンシング精度が 3 人までは性能が保たれ，4 人以降で精度が減少するという先述の考察と一貫している．従って，6 人歩行について 6 人同時に捉えられていないことを考慮すると，3 人以下の歩行は Kinect が有利であり，4 人以上の歩行は OpenPose がやや有利，7 人以上の歩行について Kinect の制約から OpenPose が有利である，と結論づける．

表 A.6　混同行列 (OpenPose)

真値＼予測	*A*	*B*	*C*	*D*	*E*	*F*	*G*	FRR(%)
A	2	0	2	1	0	0	1	66.7
B	0	2	1	0	0	3	0	66.7
C	0	0	3	2	0	1	0	50.0
D	0	0	1	4	0	1	0	33.3
E	1	0	1	0	2	0	2	66.7
F	0	1	1	1	0	2	1	66.7
G	1	0	1	1	0	0	3	50.0
FAR(%)	5.6	2.8	19.4	13.9	0.0	13.9	8.3	-

表 A.7　混同行列 (Kinect)

真値＼予測	*A*	*B*	*C*	*D*	*E*	*F*	*G*	FRR(%)
A	5	0	1	0	0	0	0	16.7
B	0	4	1	0	0	0	1	33.3
C	1	0	4	0	0	1	0	33.3
D	0	0	0	5	0	0	1	16.7
E	0	0	0	0	5	0	1	16.7
F	0	0	1	0	0	4	0	16.7
G	0	0	0	2	2	0	2	66.7
FAR(%)	2.8	0.0	8.3	5.6	5.6	2.8	8.3	-

A.5　結論

Kinect のセンシング精度は Kinect に対してまっすぐ歩行するとき高く，そうでないとき精度が低下する．また，3 人までは精度を落とさず姿勢推定が出来るが，4 人以上で精度が低下する．個人識別については，3 人までは Kinect が有利であり，4 人以上では OpenPose がやや有利であることを実験に基づき明らかにした．今後の課題として，歩行実験の参加者数を増やすことや，7 人以上の歩行に対する個人識別について調べることが挙げられる．

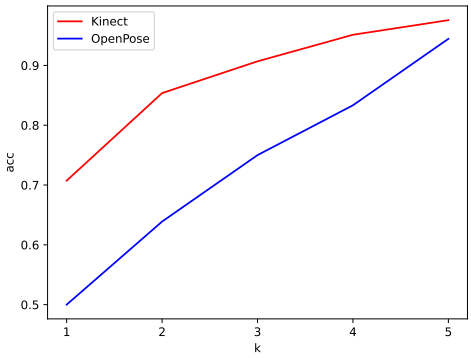

図 A.12　単独歩行 top_kAcc

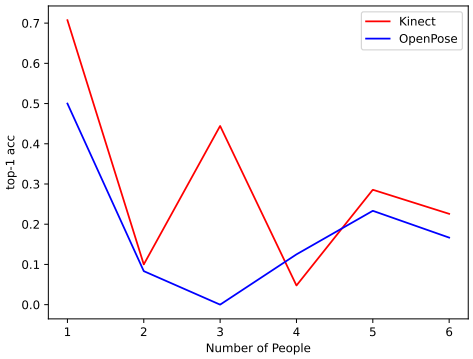

図 A.13　人数変化に伴う精度推移 (top-1 acc)

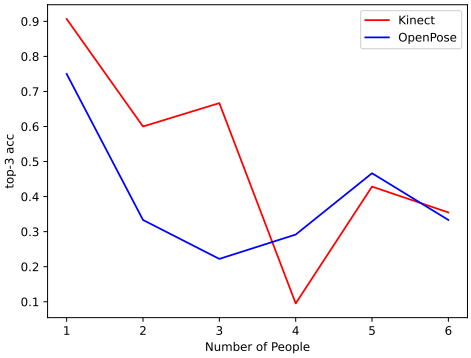

図 A.14　人数変化に伴う精度推移 (top-3 acc)

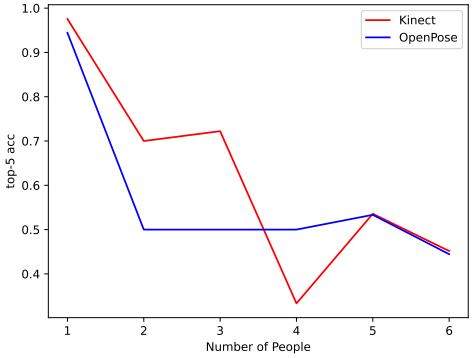

図 A.15　人数変化に伴う精度推移 (top-5 acc)

参考文献

[1] 森 駿文，菊池 浩明，“歩容データの DTW 距離に基づく個人識別手法の提案と外乱に対する評価”，情報処理学会論文誌，Vol.60, No.9, 1538-1549, 2019.

[2] Ju Han, Bir Bhanu，“Individual recognition using gait energy image”, IEEE transactions on pattern analysis and machine intelligence, 28(2)，316-322，2005.

[3] Zhe， Gines，Tomas，Shih-En，Yaser，“OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”，CVPR，pp.7291-7299，2017.

[4] Julieta, Rayat, Javier, James，“A simple yet effective baseline for 3d human pose estimation”, ICCV， pp. 2640-2649, 2017.

[5] Thomas Sanchez Lengeling，“Kinect v2 library for Processing”(`https://github.com/ThomasLengeling/KinectPV2`)，2016.

[6] 渡邊宏，“「Kinect v2」はここがスゴい！新旧比較と Kinect による NUI 開発の最前線”，MONOist，2014.

[7] Sakoe, H. and Chiba, S，“Dynamic Programming Algorithm Optimization for Spoken Word Recognition, IEEE Transaction on Acoustics, Speech, and Signal Processing”,Vol.ASSP-26, No.1, pp.43–49，1978.

[8] 三好駿, 森駿文, 菊池浩明，“歩容データからの属性暴露リスクについて”，情報処理学会第 81 回全国大会, pp.3_421-3_422, 2019.

[9] 阪田 篤哉, 武村 紀子, 八木 康史，“多段階畳み込みニューラルネットワークを用いた歩容に基づく年齢推定”， 2018 年 5 月コンピュータビジョンとイメージメディア研究会，吹田，Vol. 2018-CVIM-212，No. 23，pp. 1-5，May 2018.

付録 B

分担表

作業の分担を次の表 B.1 に示す.

表 B.1　作業分担表

作業	當麻	谷口
歩行実験	実験プログラムの作成 データの集計　など	実験参加同意書の作成 実験場所の確保　など
データの前処理	Kinect で取得した歩容データの整形	Kinect で取得した映像から OpenPose を用いて 歩容データを取得，整形
個人識別	Kinect で取得したデータに対する 識別プログラムの実装	OpenPose で取得したデータに対する 識別プログラムの実装
センシング誤り評価	Kinect のセンシング誤りについて実験	—