

Title	音響特徴とパーソナリティーおよび年齢との関連に関する研究
Author(s)	LYU, Ang
Citation	
Issue Date	2021-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/17140
Rights	
Description	Supervisor: 党 建武, 先端科学技術研究科, 修士 (情報科学)

修士論文

音響特徴とパーソナリティおよび年齢との関連に関する研究

LYU ANG

主指導教員 党 建武

北陸先端科学技術大学院大学
先端科学技術研究科
(情報科学)

令和2年3月

Abstract

The education and development of personality are an important issue concerned by whole society. For schools, the development of students' personality is a topic of continuous concern. For companies, it is also extremely important to find employees with suitable personalities. Traditional paper-based testing methods rely on the cooperation of the experimenter and consume a lot of time. So in recent years, the research of automatically predicting personality has attracted researchers' attention.

In previous studies, researchers used methods such as machine learning and deep learning to predict personality through acoustic features and explored the contribution of acoustic features to personality prediction. However, a person's acoustic features and personality will change with age. Therefore, it is necessary to consider the change of age by using acoustic features to predict personality. But, the age distribution of experimenters was not considered when predicting personality in previous studies. In particular, adolescents are the main period for personality developing. At the same time, their acoustic features also change with growing.

Therefore, in this study, we constructed a personality classification model to study the contribution of acoustic features to the personality classification of adolescents at various ages. We extracted 43 elementary school students, 40 junior high school students, 36 high school students, and 40 college students, a total of 159 people's dialogue acoustic features and tested their personality. Then, through machine learning, different acoustic feature sets are used to train the personality classification model to verify the contribution of acoustic features to the personality of each age group.

The experimental results show that each acoustic feature has a different contribution to the personality model factor classification, and with the change of age, the contribution of the acoustic feature will also change. Fundamental frequency contributes to the openness of personality classification, but it has no special correlation with age. Mel-frequency cepstral coefficients contribute to the agreeableness of personality classification, but like the fundamental frequency has no special correlation with the age. Root mean square of energy contributes to the classification of openness and agreeableness. It has a significant effect on the openness of adults or people close to adult age and the coordination of the classification of people under high school students. The zero-crossing rate has a significant contribution to the openness classification of adults or people close to adult age. The harmonics-to-noise ratio contributes to the agreeableness classification of people of all ages. And, as the age grows, the contribution of the harmonics-to-noise ratio to openness classification will become smaller.

Based on our study's result, different acoustic features have different contributions to personality classification, and the degree of contribution is also related to age. At the same time, we found that the younger the experimenter has lower accuracy of personality classification. We think that in childhood, children's vocal cords and vocal organs are still developing, and there are uncertain factors such as language ability and acoustic characteristics. We need to combine features other than acoustic features to classify the personality of adolescents.

Keywords: Personality, Acoustic characteristics, Adolescent, Mechanical learning.

目次

第1章 はじめに	1
1.1 研究背景.....	1
1.2 研究目的.....	2
1.3 本論文の構成	2
第2章 関連研究	4
2.1 パーソナリティに関する研究	4
2.2 音響特徴とパーソナリティの関連性に関する研究	6
2.3 音響特徴によるパーソナリティ分類に関する研究	7
2.4 年齢とパーソナリティの関連性に関する研究	8
2.5 年齢と音響特徴の関連性に関する研究	9
第3章 研究方法	10
3.1 問題点及び研究方法.....	10
3.2 データベース	10
3.2.1 対話録音データ	11
3.2.2 パーソナリティデータ.....	12

3.3 音響特徴抽出	14
3.4 パーソナリティデータ	15
第4章 実験	16
4.1 Extreme learning machine	16
4.2 訓練方法	18
第5章 実験結果	20
5.1 すべての音響特徴でモデルを訓練する	20
5.2 F0を取り除く	22
5.3 MFCCを取り除く	24
5.4 RMS Energyを取り除く	27
5.5 ZCRを取り除く	30
5.6 HNRを取り除く	33
第6章 考察	36
第7章 おわりに	40

目次

図 3.2.1.1 : HSK 中国語検定の口試受験の問題	11
図 3.2.1.2 : 録音データのノイズの例	12
図 3.2.1.3 : 調査対象のみの録音データを切り出す	12
図 3.2.2.1 : Neo FFI 人格検査	1 错误!未定义书签。
未找到图形项目表。图 5.2.1 : 各年齢層で全特徴と F0 を取り除く結果の比較	2 错误!未定义书签。
未找到图形项目表。图 5.4.1 : 各年齢層で全特徴と RMS Energy を取り除く結果の比較	28
图 5.5.1 : 各年齢層で全特徴と ZCR を取り除く結果の比較	31
图 5.6.1 : 各年齢層で全特徴と HNR を取り除く結果の比較	34

表目次

表 3.3.1: 抽出した音響特徴	1 错误!未定义书签。
未找到图形项目表。表 5.1.1: すべての音響特徴の場合でモデルの分類精度	20
表 5.2.1: F0 を取り除く場合でモデルの分類精度	22
表 5.2.2: すべての特徴と F0 を取り除く場合での差の t 検定	23
表 5.3.1: MFCC を取り除く場合でモデルの分類精度..2	错误!未定义书签。
未找到图形项目表。表 5.4.1: RMS Energy を取り除く場合でモデルの分類精度	27
表 5.4.2: すべての特徴と RMS Energy を取り除く場合での差の t 検定..	28
表 5.5.1: ZCR を取り除く場合でモデルの分類精度	30
表 5.5.2: すべての特徴と ZCR を取り除く場合での差の t 検定	31
表 5.6.1: HNR を取り除く場合でモデルの分類精度	33
表 5.6.2: すべての特徴と HNR を取り除く場合での差の t 検定	34

第1章 はじめに

1.1 研究背景

パーソナリティの育成と発達、常に社会に注目されている課題である。教育分野では、学生のパーソナリティの育成がますます注目されている。学生が健康に成長できるには、学校が学生のパーソナリティを把握する必要がある。また、企業で人材を採用する場合、協力が得意、コミュニケーション能力が高い人が高く評価されることが多い。適切な社員を探すために、求職者のパーソナリティも評価基準の1つである。これらの必要を満たすために、アンケートを通じてパーソナリティを測定する方法が発明された。従来のアンケート測定方法の効果は、被験者の協力の程度に依頼し、大量な時間もかかる。それゆえに、高効率で自動的なパーソナリティ分類の研究が注目されている。その中、機械学習と深層学習の手法で、音響特徴によるパーソナリティ分類する研究があった。

しかし、人の音響特徴とパーソナリティーは年齢によって変化している。音響特徴を通して、パーソナリティを分類するには、年齢の変化を考える必要があるが、今までのパーソナリティ分類モデルには年齢を考慮していなかった。特に青少年期に対して年齢に連れてかなり変化しており、パーソナリティーは年齢により、変化していることが考えられる。

1.2 研究目的

青少年期は人間の育成と成長の重要な期間である。青少年のパーソナリティを把握することで、青少年の教育と育成に役に立つことができる。人の音響特徴とパーソナリティは年齢によって変化している。しかし、今までの音響特徴によるパーソナリティを分類するモデルは、人の年齢層を考えていない。特に、青少年期がパーソナリティ特性の変化の主要な期間であるので、パーソナリティを分類するには、年齢因子を考慮する必要があると想定している。

したがって、本研究の目的は青少年に着目し、音響特徴を用いてパーソナリティと年齢の関連性を明らかにすることである。

1.3 本論文の構成

本論文は、7章で構成されている。

第1章ははじめに、本論文の研究背景と目的を紹介する。

第2章は関連研究である。まず、パーソナリティに関する研究を紹介する。次に、音響特徴とパーソナリティの関連性に関する研究と音響特徴によるパーソナリティ分類に関する研究を紹介する。最後に年齢とパーソナリティ及び音響特徴の関連性に関する研究を紹介する。

第3章は研究方法を説明する。本研究で利用したデータベース、音響特徴集及び機械学習モデルを紹介する。

第 4 章は実験である。ここで、機械学習の手法で音響特徴によるパーソナリティ分類の方法を説明する。

第 5 章は実験結果である。ここで、パーソナリティ分類の分析結果を説明する。

第 6 章は考察である。パーソナリティ分類の結果に基づき、異なる年齢層で音響特徴とパーソナリティ分類の関連性を説明する。

第 7 章のおわりにでは、実験の結果に基づき、分析した結論をまとめる。最後に、本研究の不足と未来の研究方向を説明する。

第2章 関連研究

2.1 パーソナリティに関する研究

1990年代以降、パーソナリティ心理学の領域で Big Five モデルが飛躍的に発展した。Big Five モデルとは、人のパーソナリティを幅広い観点から捉えた場合、評価の次元として安定した5つの特徴が見出されるものである[1]。その特性は：

E：外向性（Extroversion）

N：精神的安定性（Neuroticism）

O：経験への開放性（Openness to experience）

C：勤勉性（Conscientiousness）

A：協調性（Agreeableness）

外向性とは、対人関係の量と密度、刺激の必要性、そして喜びを得る能力である。この次元では、社会的で活動的な人と、真面目で、恥ずかしがり屋で、静かな人とを対比する。この次元は、対人関係のレベルと活力のレベルという2つの性質によって測定できる。

精神的安定性は、人の感情調節のプロセスを反映し、否定的な感情や感情的な不安定さを経験する個人の傾向を反映している。精神的不安定性が高い人は、

非現実的な考え、過度の要求や衝動を抱く傾向があり、怒り、不安、うつ病などの否定的な感情を持つ可能性が高くなる。彼らは一般の人々よりも外部の刺激に強く反応し、感情的な調節と対処能力が低く、しばしば悪い感情状態にある。そして、これらの人々は、思考、意思決定、および外部からの圧力に効果的に対処する能力が不十分である。

経験への開放性は、自分自身の経験を理解するための積極的な探求、及び不慣れな状況の寛容と探求の能力である。この次元は、好奇心が強く、非伝統的で創造的な人と、伝統的で、芸術に興味がなく、分析できない人を比較する。開放性が高い人は抽象的な思考を好み、幅広い関心を持っている。開放性が低い人はより伝統的である。

勤勉性は、私たちが自分の衝動を制御、管理、規制する能力を指す。この次元は目標に向ける行動する人の持続性、動機を評価する。つまり、信頼できる、要求の厳しい人と怠惰な人を比較する。また、人の自己管理の能力を反映している。

協調性は、他人に対する態度を評価する。これらの態度には、一方では、親密で、信頼でき、寛大で、心の優しいものが含まれる。他方では、敵対的、容赦のない、復讐心、容赦ないなどが含まれる。協調性が高い人は、ほかの人の利益のために自分の利益を放棄する思い方がある。協調性が低い人は、自分の利益をほかの人の利益より優先する思い方がある。

このモデルは大量なパーソナリティ表現語を基礎となり、多変量解析法の発展とともに5つのパーソナリティ特性に収束されたものである。異なる言語でも同じ次元があると確認された。

パーソナリティは、人の意識、感情と行動に関係し、心理健康、人間関係に影響している。特に青少年は身体的と精神的に成長するので、学童期と若年成人期はパーソナリティの育成の重要な期間である。パーソナリティの育成を研究するには、パーソナリティの育成と教育の基礎になる[2, 3, 4, 5, 6, 7]。

2.2 音響特徴とパーソナリティの関連性に関する研究

音響特徴によるパーソナリティを分類する研究は、まず両者の関係を明らかにすることが重要である。したがって、音響特徴とパーソナリティの関連性に関する研究を整理する。

内田の研究では一部の音響特徴が Big Five パーソナリティモデルに基づいたパーソナリティスコアに反映できることを明らかにした。この研究では、発話速度、音声の基本周波数、音声の抑揚に着目し、発話対象の性格印象について分析した。その結果、音声の韻律的特徴量と被験者のパーソナリティの間に非線形な関連性を明らかにした[8]。

Donna Erickson の研究では、他人の人格判断に影響を与えるいくつかの音響的特徴が紹介した。この研究では、F0、エネルギー、ノイズ比率、フォルマント

などの音響特性を使用してパーソナリティを分類した。その結果、声の変化の範囲が最も広い人々は、外向性と勤勉性が高い。ノイズ比率が高い人は、協調性が低い[9]。

2.3 音響特徴によるパーソナリティ分類に関する研究

Gelareh Mohammadi の研究では、音響特徴によるパーソナリティ予測モデルを構築した。彼らは、6つの音響特徴と4つの統計特徴を使用して、パーソナリティを自動的に予測する方法を提出した。その結果、Big Five モデルのそれぞれのパーソナリティスコアを高精度（特定の特徴に応じて 70%以上）で予測できた[10]。

Tim Polzehl の研究では、音響特徴によるパーソナリティの分類モデルを構築された。分類の結果について、精神的安定性と外向性の2つの次元に対する分類の効果が特に優れる[11]。

Chao S. Hu の研究は、異なる感情状態で話者の性格特性と発話周波数スペクトルの相関関係を分析した。さまざまな感情状態での反応を記録します。彼らは音素の最初の3つのフォーマントの周波数に着目し、パーソナリティを分類した。結果は、実験者が通常感情状態からポジティブな感情状態へ変化するとき、話者の感情の安定性と社会的勇気に大きく関係している。関連する感情状態では、音声スペクトルパラメータは、話者の性格特性に大きく関係している。

さらに、異なる感情状態と話者の反応の間の相関も異なる。この研究の結果は、音声を通して話者のパーソナリティを検出できることを示している[12]。

これらの研究では、音響特徴によるパーソナリティを分類する可能性を示したが、分類の精度を向上させる必要がある。

2.4 年齢とパーソナリティの関連性に関する研究

心理学の研究では、思春期と若年成人期が、パーソナリティ特性の形成する主要な期間であり、青少年の声も変化していることを議論した。結果は、青春期の子供の外向性と開放性が一般的に高くなる。年齢とともに、人々の自覚性と感情の安定性は向上することを示している[13, 14]。

これらの研究は、子供と青少年のパーソナリティ形成の変化傾向を示している。研究結果によると、子供の外向性は学童期から青年期まで減少し続けている。そして、協調性は学童期初期から上昇し続け、学童期後期から青年期初期に徐々に低下し、青年期中期および後期に再び上昇する。勤勉性の発達傾向は協調性と同じである。精神的安定性の発達と変化は性別に関係している。女の子の精神的安定性は、学童初期に低いが、学童期後期から青年期中期に上昇し続け、青年期後期から低下する。男の子の精神的安定性は、学童期から青年期後期までに低下し続ける。開放性は学童期からわずかに上昇し、学童期後期から青年期中期までに低下し、思春期後期に上昇し続けた[15, 16, 17, 18, 19, 20]。

これらの研究では、パーソナリティは年齢とともに変化することを明らかにしたが、今までの音響特徴によるパーソナリティを分類するモデルは、年齢因子を考えていなかった。本研究は年齢因子を追加し、異なる年齢で音響特徴がパーソナリティ分類への貢献を分析する。

2.5 年齢と音響特徴の関連性に関する研究

音声言語医学領域の研究は青少年群及び中年群と老年群の基本周波数の差を分析した。成人女性の話声位 (Speaking Fundamental Frequency) について得られた興味深い知見として、20代の人と比較して30代と40代の人と話声位が明らかに低下していることをわかった。異なる年齢の男性と女性の話声位の変化も異なる[21]。

男性話者について、10年から24年にわたって収録した録音から抽出した5種類の音響特徴を加齢との関係について分析した。個人差が大きいこと、F0が低下する傾向がある話者が多いこと、高域の喉頭雑音量は加齢によって増加する傾向があることなどがわかった[22]。

これらの研究では、音響特徴は年齢とともに変化することを明らかにした。そのゆえ、異なる年齢層で変化している音響特徴はパーソナリティ分類への貢献も変化している可能性がある。その貢献を分析することで、パーソナリティ分類に有用な知見を提供する可能性があると考えられている。

第3章 研究方法

3.1 問題点及び研究方法

先行研究においては、機械学習、深層学習などの手法を用いて、音響特徴によるパーソナリティを分類するモデルが提案されてきたが、年齢を考慮しなかった。しかし、人の音響特徴は、特に青少年期に対して年齢とともにかなり変化しており、パーソナリティーは年齢とともに変化している。したがって、音響特徴を用いて、青少年のパーソナリティを分類するには、年齢の要素を考えるべきである。

本研究では、機械学習を用いて、異なる音響特徴を利用し、小・中・高と大学の学生のパーソナリティを分類し、音響特徴とパーソナリティ及び年齢の関連を分析した。

3.2 データベース

本研究では、Zhai の「パーソナリティと音声コミュニケーション能力との関連についての研究」で提出されたデータベースを利用した。このデータベースには、中国人の被験者 159 名（小学生 43 人、中学生 40 人、高校生 36 人、大学生 40 人）の対話録音（一人 3 分ごと）と NEO FFI 人格検査結果が含まれている。159 人の被験者は中国中部地域にある四つの学校（スチールニュービレッジ

小学校, 合肥工科大学の所属中学校、高校, 合肥師範大学) の学生である [23]。

3.2.1 対話録音データ

HSK中国語レベル試験問題

- 1、簡単な自己紹介をしてください。
- 2、一番楽しいことは何ですか？
- 3、一番住みたいところはどこですか？理想的な生活環境は何ですか？
- 4、困難にあったとき、誰と相談して、問題を解決しますか？どうして？
- 5、どんなテレビ番組が好きですか？どうして？
- 6、一番悲しいことは何ですか？

図 3.2.1.1 : HSK 中国語検定の口試受験の問題

録音内容は HSK 中国語検定の口試受験の問題の答えである。HSK 中国語検定の口試は受験生の会話能力を判定する試験である [24] (図 3.2.1.1)。音データは、インタビューの形で六つの答えである。録音環境について、質問者と調査対象が同じ部屋に面と向かって対話する。録音設備は調査対象との距離が変わらない。録音設備は SONY ICD-UX560F レコーダーである。

元の録音データでは、録音にノイズが含まれていて、録音は質問者と調査対象、二人の音声と一緒に含まれているため、私達は録音データを前処理した。ここで、Audacity という無料なデジタル音声編集ソフトウェアを利用した。まず、録音

データのノイズを取り除いた（図 3.2.2.2）。次に、調査対象の録音のみ切り出した（図 3.2.1.3）。

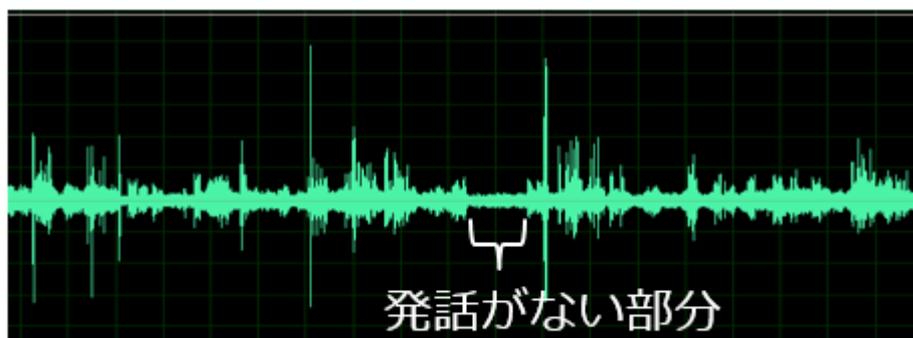


図 3.2.1.2：録音データのノイズの例

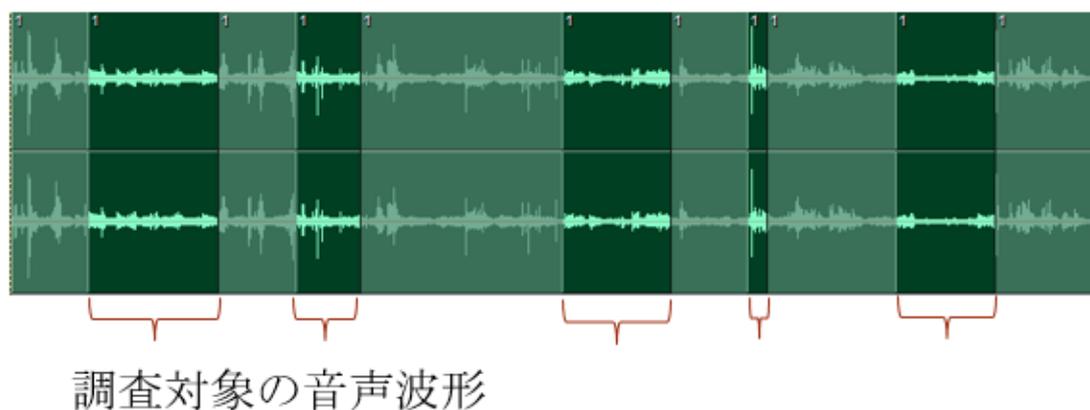


図 3.2.1.3：調査対象のみの録音データを切り出す

3.2.2 パーソナリティデータ

本研究で使したパーソナリティデータは各被験者の NEO FFI 人格検査結果である。（図 3.2.2.1）

Neo FFI 人格検査 アンケート

以下の質問をお読みください。各質問には5つの選択肢があります。説明が明らかにあなたに合わない場合、またはあなたが強く同意しない場合は、1を選択してください。ほとんどの場合で説明があなたに合わない場合は、2を選択してください。説明が半分正しく、半分間違っている場合は、3を選択してください。ほぼあなたと一致しているか、あなたが同意している場合は、4を選択してください。説明が明らかにあなたと一致している場合、またはあなたが非常に同意している場合は、5を選択してください。

問題	当てはまらない	どちらかというと当てはまらない	どちらとも言えない	どちらかというと当てはまる	当てはまる
1、うつ病になりやすい人。	1	2	3	4	5
2、周りにたくさんの友達がいるのが好き。	1	2	3	4	5
3、ファンタジーに没頭し、実現できるすべてのことを探求するのが好き。	1	2	3	4	5
4、出会うすべての人に礼儀正しく対する。	1	2	3	4	5
5、いつも持ち物を清潔に保つ。	1	2	3	4	5
6、時々私は怒りと恨みに満ちている	1	2	3	4	5
7、笑うのが好き	1	2	3	4	5
8、新しい趣味を育てていくのが好き。	1	2	3	4	5
9、時々、脅迫や世辞などの異なる方法を使用して、他の人に自分のやりたいことをするように説得する。	1	2	3	4	5

図 3.2.2.1 : Neo FFI 人格検査

Big Five 人格検査の短縮版である NEO FFI 人格検査は Big Five モデル理論に基づいており、60 の質問が含まれている。実用化の必要性のために、NEO-FFI はプロジェクト要因を分析した後に NEO-PI によって簡略化される。また、NEO-FFI は、NEO-PI-R の有効性と本質的な内容を継承している。この 2 つの検査は、関連性が高く、信頼性が高く、効果的で、世界中で広く使用されているパーソナリティ評価手法である。NEO FFI 人格検査の質問用紙は質問と回答欄

が含まれている。質問について、60 問題がある。1 から 5 まで、5 段階で回答を得る。回答欄の下に各問題の得点があり、その得点を足すと、各因子の素点が算出できる(各次元は 60 点満点で評価される)。この五つの点数に基づいてパーソナリティを判断できる [25]。

3.3 音響特徴抽出

本研究では、INTERSPEECH 2009 Emotion Challenge feature set [26] 特徴集を利用した。

音響特徴	統計特徴
ZCR	mean
RMS Energy	standard deviation
F0	kurtosis, skewness
HNR	extremes: value, rel.position, range
MFCC 1-12	linear regression: offset, slope, MSE

表 3.3.1: 抽出した音響特徴

この特徴集には、音響特徴 16 個と統計特徴 12 個が含まれている。16 個の音響特徴は基本周波数 (F0)、MFCC 係数 1-12、零交差率 (ZCR)、エネルギーの二乗平均平方根値 (RMS Energy)、ノイズ比率 (HNR) 。16 個の統計特徴は平均値、標準偏差、尖度、歪度、極値、相対位置の極値、範囲の極値、offset の線形回帰、slope の線形回帰、MSE の線形回帰。

Opensmile という音響特徴抽出ソフトウェアを用いて、159 人の録音データ

の音響特徴量を抽出した。

3.4 パーソナリティデータ

NEO FFI 人格検査のパーソナリティスコアの満点は 60 点。NEO FFI 人格検査結果分析と分析データのバランスにより、パーソナリティレベルを 3 つに分けた。人数分布を表 3.4.1 に示す。

五因子	35 点以下	35～45 点	45～60 点
精神的安定性	101	47	11
外向性	29	64	66
開放性	18	81	60
協調性	16	78	65
勤勉性	24	54	81

表 3.4.1: パーソナリティ五因子の人数分布

第4章 実験

前述した音響特徴を使用し、機械学習手法である Extreme Learning Machines (ELM) を用いて、パーソナリティを分類し、分類精度を求めた。また、音響特徴がパーソナリティ及び年齢との関連性を分析するために、関連にある音響特徴量を1つずつ取り除き、残った特徴でモデルを訓練し、年齢層を分けて、分類精度を分析した。

4.1 Extreme learning machine

本研究では、Extreme learning machine を使い、パーソナリティの分類実験を行う。

Extreme learning machine (ELM) は、1層もしくは複数の隠れ層を有する順伝播型ニューラルネットワークである[27, 28]。類と回帰タスクを解決できる。

ELM モデルの学習手順として、以下になる：

- 1、入力層から隠れ層への重みをランダムに生成する。
- 2、活性化関数を通した隠れ層の出力を計算する。
- 3、隠れ層の出力とデータの回帰係数を求め、隠れ層から出力層への重みとする。

$$\hat{Y} = W_2 \sigma(W_1 x)$$

$$W_2 = \sigma(W_1 X) + Y$$

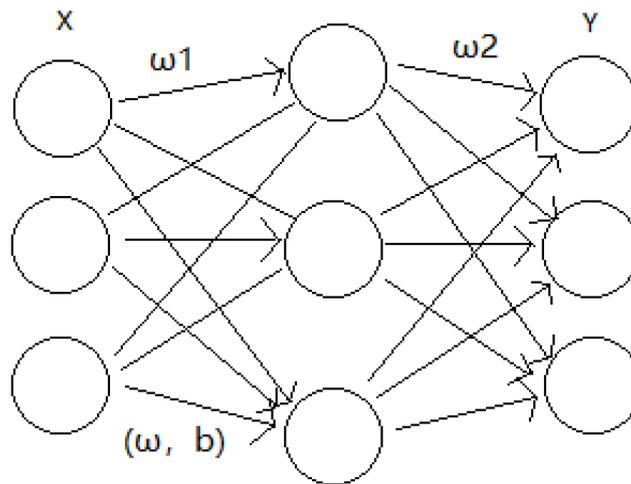


図 4.1.1 : ELM の構造

ELM モデルにはいくつかのメリットがある：

- 1、学習の速度が早い。
- 2、過学習を回避しやすい。
- 3、多様な活性化関数を利用できる。
- 4、決定するパラメータが少ない（必要なパラメータは隠れ層のノード数だけ）。

また、ELM モデルには、入力層から隠れ層への重みはランダムに 1 回決められ、アルゴリズムの実行中に調整する必要がないという特徴がある。[28]

今回の実験では、ELM モデルの隠れ層ノードの数を 1 から 30、50、100、150、200、250、300 に設定し、パーソナリティ分類の表現を統計した。最後に、隠れ層ノード数が 7 の倍数であるとき、パーソナリティ分類モデルの結果がより良くなった。そのため、以上の場合で一番高い分類精度である 7 を隠れ層ノード数として実験を行った。

4.2 訓練方法

本研究で使用したデータベースのデータ量が少ないため、K-分割交差検証 (K-fold cross-validation) を利用して、ELM モデルを訓練する。

K-分割交差検証では、データを K 個に分割する。そして、そのうちの 1 つをテスト事例とし、残る $K - 1$ 個をテストデータとする。そして、K 個に分割されたデータそれぞれをテストセットとして k 回検証を行う。そうやって得られた k 回の結果を平均して最終結果とする。

K-分割交差検証は 2 つのメリットがある：

- 1、すべてのデータはテストされているため、テストセットとトレーニングセットを分割する方法の影響を受けない。

- 2、データ量が足りない場合では、すべてのデータがトレーニングデータとして使われる。

本研究では、K-分割交差検証方法の K を 20 とする。すなわち、データを 20

回訓練する。毎回訓練するとき、95%のデータをトレーニングデータとして使用される。残った5%のデータをテストデータとして使用される。

第5章 実験結果

本章では、まずすべての音響特徴集を用い、ELM モデルを訓練し、パーソナリティの分類結果を示す。次に、関連の音響特徴を1つずつ取り除き、残った音響特徴を用い、ELM モデルに再学習させ、結果を示す。

5.1 すべての音響特徴でモデルを訓練する

すべての音響特徴を使用し、ELM モデルを訓練し、パーソナリティを分類した精度を表 5.1.1 に示す。

All feature	N	E	O	A	C
小学生	0.4168	0.3488	0.4419	0.4186	0.6279
中学生	0.45	0.375	0.6	0.5	0.55
高校生	0.444	0.5	0.6111	0.6944	0.4167
大学生	0.3	0.725	0.825	0.8	0.35

表 5.1.1: すべての音響特徴の場合でモデルの分類精度

本実験の結果は先行研究と比較すると、パーソナリティ 5 因子の精神的安定性の分類結果と中学生以下の分類結果は低い、大学生の外向性、開放性と協調性の分類結果はより優れている。また、年齢層を分けて分析すると、以下のこと

がわかった：

1、中学生の精神的安定性の正確率が一番高い、大学生の精神的安定性の正確率が一番低い。

2、年齢が上がるとともに、外向性の分類正確率が徐々に向上した。

3、年齢が上がるとともに、開放性の分類正確率が徐々に向上したが、中学生と高校生の分類正確率が同じい。

4、年齢が上がるとともに、勤勉性の分類正確率が徐々に向上した。

5、年齢が上がるとともに、協調性の分類正確率が徐々に低下した。

これは、年齢は音響特徴でパーソナリティを分類するには影響があることを証明した。

次に、単一の音響特徴がパーソナリティ分類への貢献を分析するために、関連の音声特徴を1つずつ取り除いて、異なる音響特徴データ集を用い、再学習した結果を分析した。

5.2 F0 を取り除く

F0 を 取り除く	N	E	O	A	C
小学生	0.4651	0.3721	0.3953	0.4286	0.6047
中学生	0.5	0.55	0.625	0.55	0.475
高校生	0.3333	0.6389	0.5556	0.5833	0.3889
大学生	0.3	0.625	0.675	0.75	0.35

表 5.2.1: F0 を取り除く場合でモデルの分類精度

基本周波数 (F0) を音響特徴集から取り除き、ELM モデルを訓練した。その結果を表 5.2.1 に示す。

年齢層を分けて、すべての音響特徴の場合での結果と比較した (図 5.2.1)。t 検定した結果を表 5.2.2 に示す。

F0 を取り除く場合では、以下のことがわかった：

1、勤勉性 t-検定の p 値は 0.046 (<0.05)。F0 を取り除く場合で勤勉性分類の正確率の変化は有意差がある。F0 は勤勉性の分類に貢献している。

2、中学生以下の被験者の精神的安定性の正確率が向上し、高校生になると、

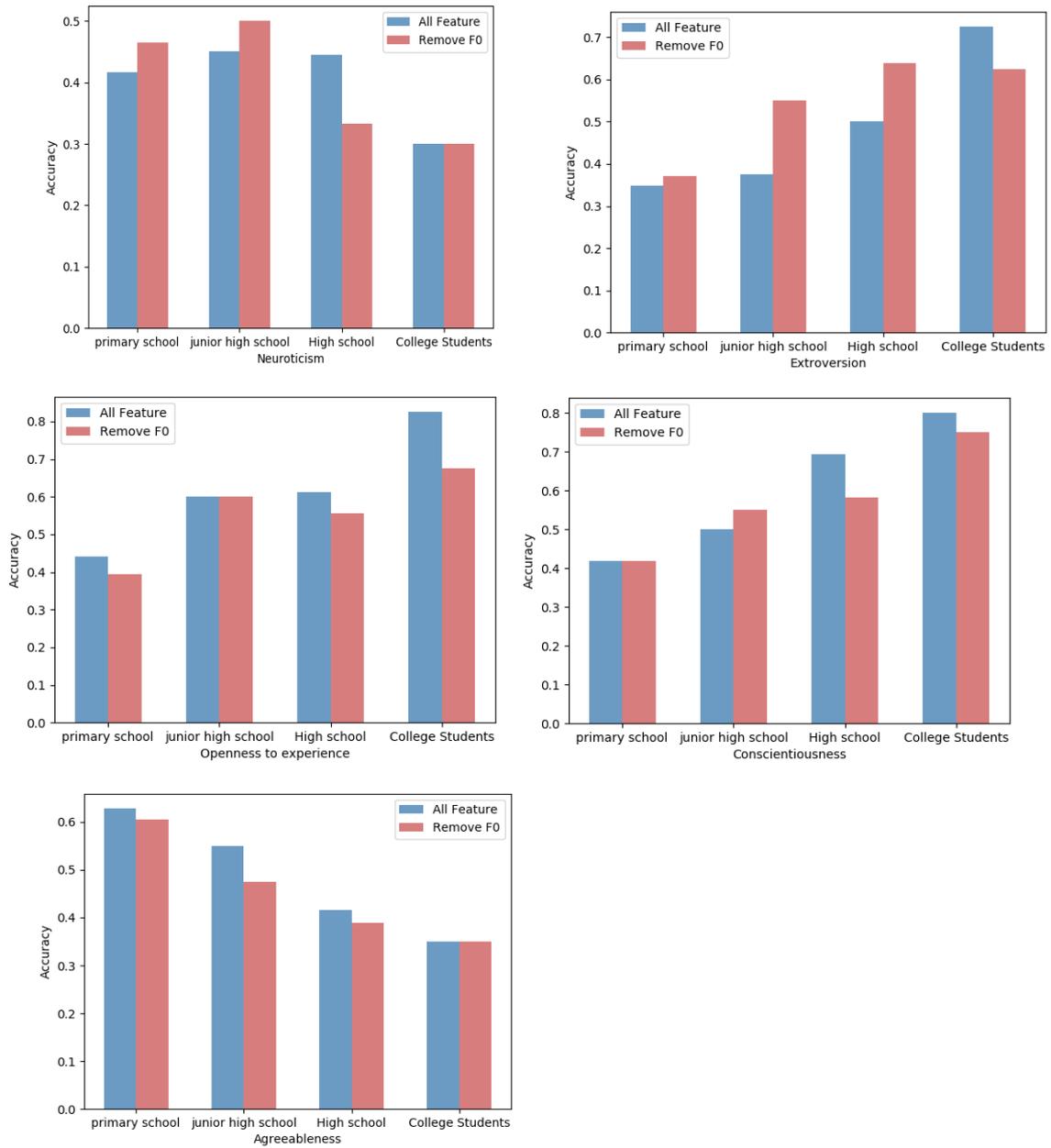


図 5.2.1：各年齢層で全特徴と F0 を取り除く結果の比較

属性	N	E	O	A	C
P 値	0.940	0.411	0.212	0.525	0.046

表 5.2.2: すべての特徴と F0 を取り除く場合での差の t 検定

精神的安定性の正確率が低下した傾向がある。

3、高校生以下の外向性の正確率が向上したが、大学生になったら逆に正確率が低下した。

4、すべての年齢層の被験者の開放性の分類正確率が低下した。

5、年齢が上がると、中学生以下の協調性の正確率が向上したが、高校生以上になったら逆に低下した。

6、すべての音響特徴の結果と異なり、中学生以下の精神的安定性の正確率が高く、高校生以上の精神的安定性が低くなった。

5.3 MFCC を取り除く

MFCC を取り除く	N	E	O	A	C
小学生	0.4419	0.3488	0.4419	0.3953	0.7209
中学生	0.55	0.475	0.6	0.475	0.525
高校生	0.3333	0.5	0.6111	0.6389	0.3056
大学生	0.3	0.7	0.775	0.775	0.35

表 5.3.1: MFCC を取り除く場合でモデルの分類精度

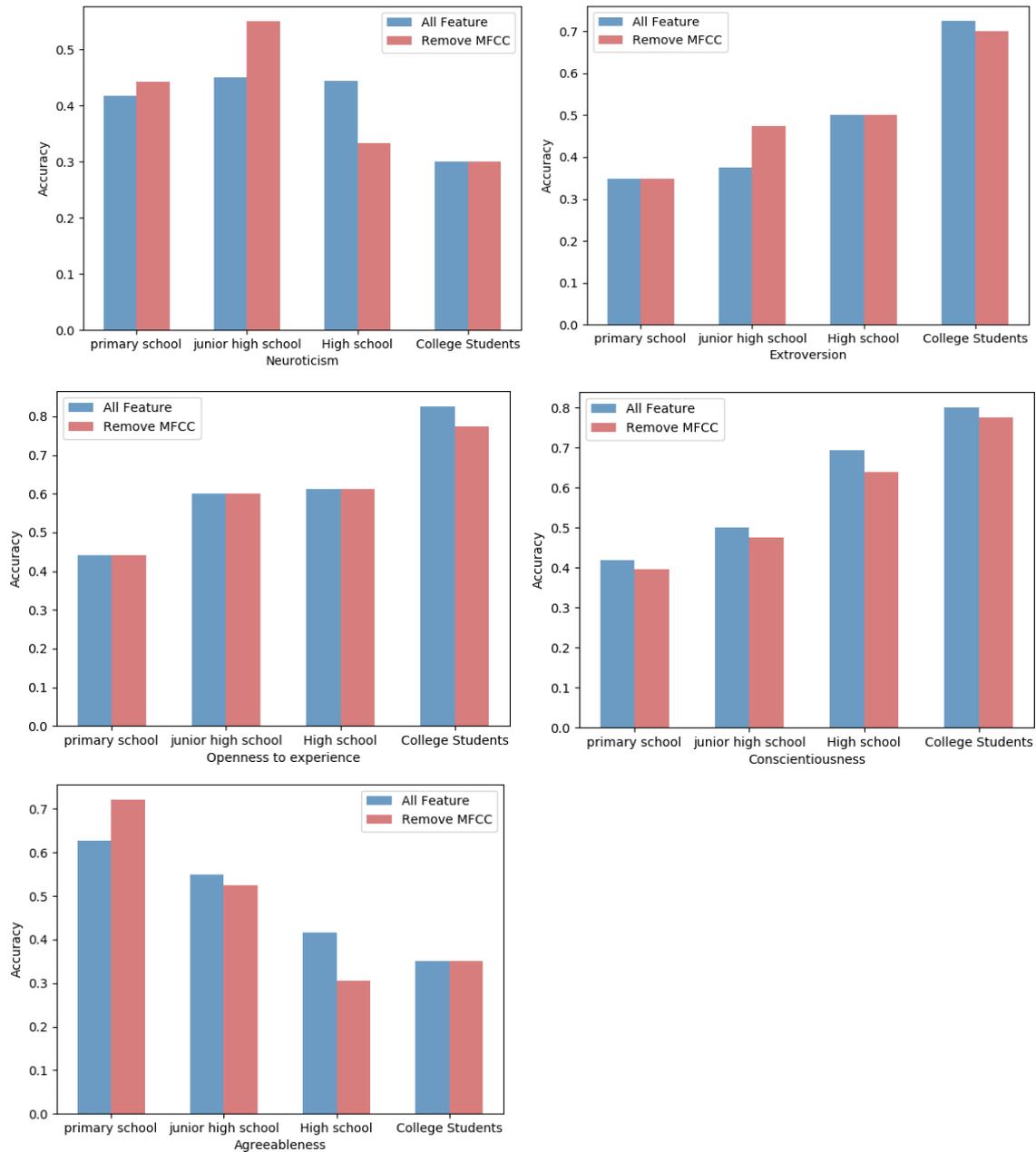


図 5.3.1 : 各年齢層で全特徴と MFCC を取り除く結果の比較

属性	N	E	O	A	C
P 値	0.939	0.547	0.391	0.026	0.814

表 5.3.2: すべての特徴と MFCC を取り除く場合での差の t 検定

MFCC を音響特徴集から取り除き、ELM モデルを訓練した。その結果を表 5.3.1 に示す。

年齢層を分けて、MFCC を取り除く場合での結果とすべての音響特徴の場合での結果を比較した (図 5.3.1)。t 検定した結果を表 5.3.2 に示す。

MFCC を取り除く場合では、以下のことがわかった：

1、協調性 t-検定の p 値は 0.026 (<0.05)。MFCC を取り除く場合で協調性分類の正確率の変化は有意差がある。MFCC は協調性の分類に貢献している。

2、中学生以下の精神的安定性の分類正確率が向上し、高校生以上の精神的安定性の分類正確率が低下した。

3、小学生と高校生の外向性の正確率が前後で変わっていないが、中学生の外向性の正確率が向上し、大学生の外向性の正確率が低下した。

4、高校生以下の開放性の正確率が変わっていないが、大学生のみの開放性の正確率が低下した。

5、小学生の勤勉性の正確率が向上したが、年齢が増えたとき、中学生と高校生の勤勉性の正確率が低下した。大学生の勤勉性の正確率が変わっていない。

5.4 RMS Energy を取り除く

RMS Energy を音響特徴集から取り除き、ELM モデルを訓練した。その結果を表 5.4.1 に示す。

RMS energy を取り除く	N	E	O	A	C
小学生	0.4884	0.3023	0.4419	0.3953	0.6047
中学生	0.45	0.4	0.6	0.425	0.575
高校生	0.5	0.5	0.5833	0.6667	0.4167
大学生	0.2	0.675	0.775	0.8	0.45

表 5.4.1: RMS Energy を取り除く場合でモデルの分類精度

年齢層を分けて、RMS Energy を取り除く場合での結果とすべての音響特徴の場合での結果を比較した (図 5.4.1)。t 検定した結果を表 5.4.2 に示す。

RMS Energy を取り除く場合では、以下のことがわかった：

1、協調性 t-検定の p 値は 0.048 (<0.05)。RMS Energy を取り除く場合で協調性分類の正確率の変化は有意差がある。RMS Energy は協調性の分類に貢献している。

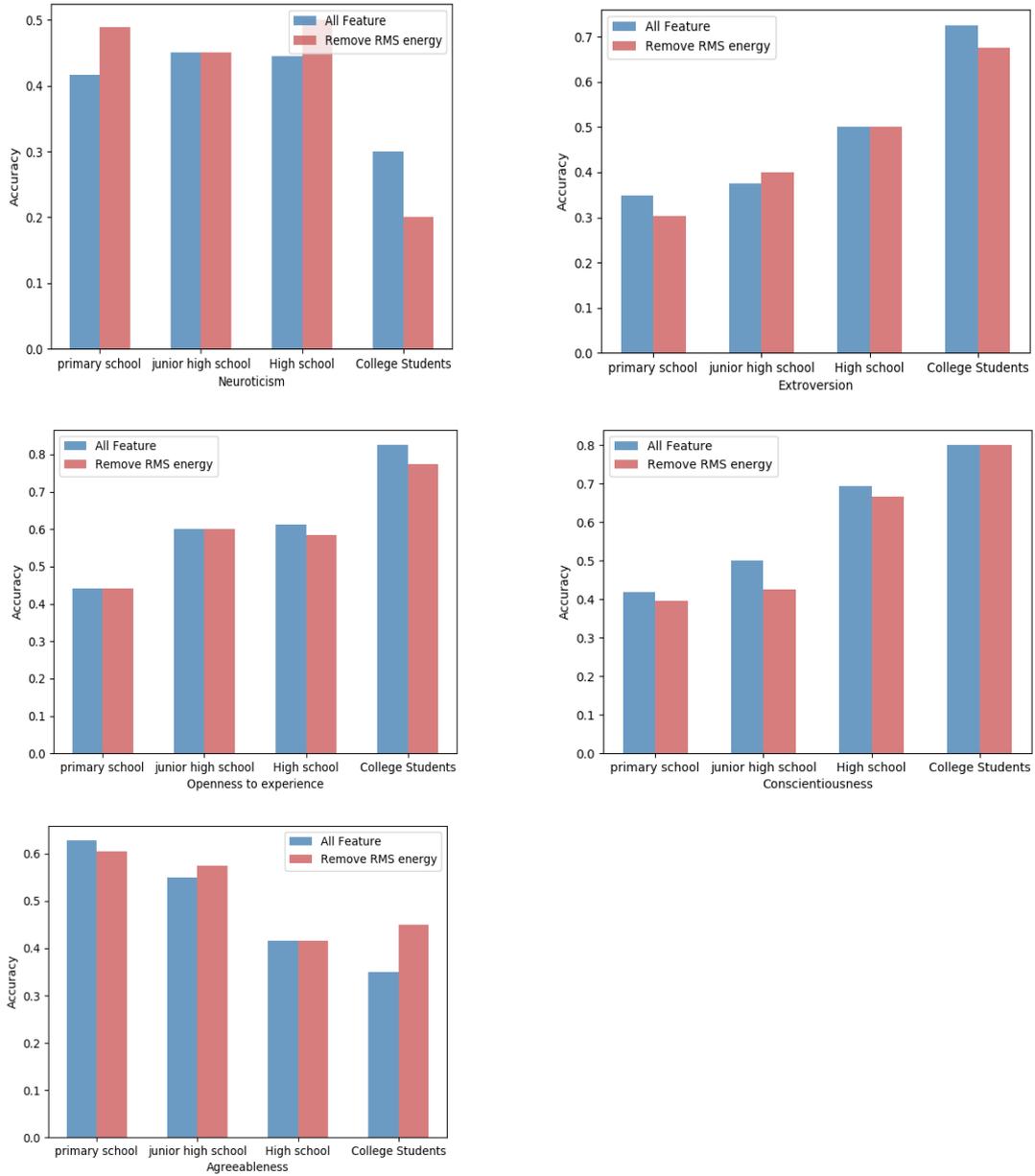


図 5.4.1 : 各年齢層で全特徴と RMS Energy を取り除く結果の比較

属性	N	E	O	A	C
P 値	0.870	0.4	0.207	0.048	0.411

表 5.4.2: すべての特徴と RMS Energy を取り除く場合での差の t 検定

2、精神的安定性について、小学生と高校生の精神的安定性の正確率向上した。中学生の精神的安定性の正確率が変わっていない。大学生の精神的安定性の正確率が低下した。

3、外向性について、小学生と大学生の外向性の正確率が低下した。中学生の外向性の正確率が向上した。高校生の外向性の正確率が変わっていない。

4、開放性について、中学生以下の開放性の正確率が変わっていない。高校生以上の開放性の正確率が低下した。

5、勤勉性について、小学生の勤勉性の正確率が低下し、中学生と大学生の勤勉性の正確率が向上した。高校生の勤勉性の正確率が変わっていない。

6、各年齢層の外向性、開放性、協調性の正確率の変化傾向は全音響特徴の結果と似ている。

5.5 ZCR を取り除く

ZCR を音響特徴集から取り除き、ELM モデルを訓練した。その結果を表 5.

5.1 に示す。

ZCR を 取り除く	N	E	O	A	C
小学生	0.4419	0.3488	0.4419	0.3953	0.6047
中学生	0.375	0.5	0.6	0.45	0.525
高校生	0.5278	0.5	0.5833	0.6667	0.4722
大学生	0.3	0.675	0.775	0.8	0.325

表 5.5.1: ZCR を取り除く場合でモデルの分類精度

年齢層を分けて、ZCR を取り除く場合での結果とすべての音響特徴の場合での結果を比較した (図 5.5.1)。t 検定した結果を表 5.5.2 に示す。

ZCR を取り除く場合では、以下のことがわかった：

1、協調性 t-検定の p 値は 0.041 (<0.05)。ZCR を取り除く場合で協調性分類の正確率の変化は有意差がある。ZCR は協調性の分類に貢献している。

2、精神的安定性について、小学生と高校生の精神的安定性の正確率が向上し、中学生の精神的安定性の正確率が低下した。大学生の精神的安定性の正確率が

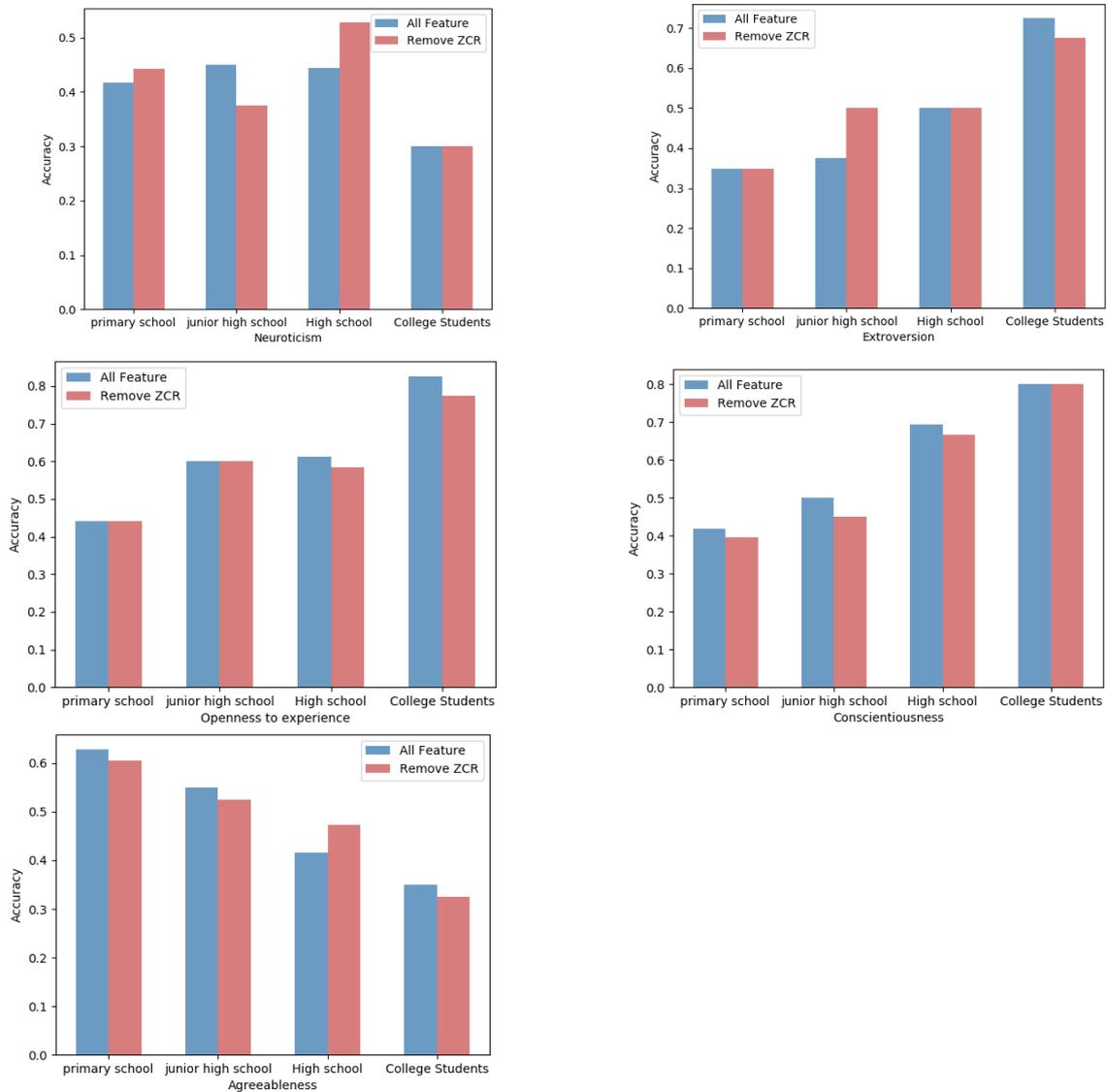


図 5.5.1 : 各年齢層で全特徴と ZCR を取り除く結果の比較

属性	N	E	O	A	C
P 値	0.813	0.650	0.207	0.041	0.839

表 5.5.2: すべての特徴と ZCR を取り除く場合での差の t 検定

変わっていない。

3、外向性について、小学生と高校生の外向性の正確率が変わっていない。中学生の外向性の正確率が向上し、大学生の外向性の正確率が低下した。

4、開放性について、小学生と中学生の開放性の正確率が変わっていない。高校生と大学生の開放性の正確率が低下した。

5、勤勉性について、小学生、中学生と大学生の勤勉性の正確率が低下し、高校生の勤勉性の正確率が向上した。

5.6 HNR を取り除く

HNR を音響特徴集から取り除き、ELM モデルを訓練した。その結果を表 5.

6.1 に示す。

HNR を取り除く	N	E	O	A	C
小学生	0.4884	0.3256	0.4419	0.3953	0.5116
中学生	0.475	0.425	0.6	0.45	0.55
高校生	0.3333	0.4722	0.6111	0.6667	0.3889
大学生	0.3	0.775	0.775	0.775	0.4

表 5.6.1: HNR を取り除く場合でモデルの分類精度

年齢層を分けて、HNR を取り除く場合での結果とすべての音響特徴の場合での結果を比較した (図 5.6.1)。t 検定した結果を表 5.6.2 に示す。

HNR を取り除く場合では、以下のことがわかった：

1、協調性 t-検定の p 値は 0.015 (<0.05)。HNR を取り除く場合で協調性分類の正確率の変化は有意差がある。HNR は協調性の分類に貢献している。

2、精神的安定性について、小学生と中学生の精神的安定性の正確率が向上し、高校生の精神的安定性の正確率がかなり低下した。大学生の精神的安定性の正

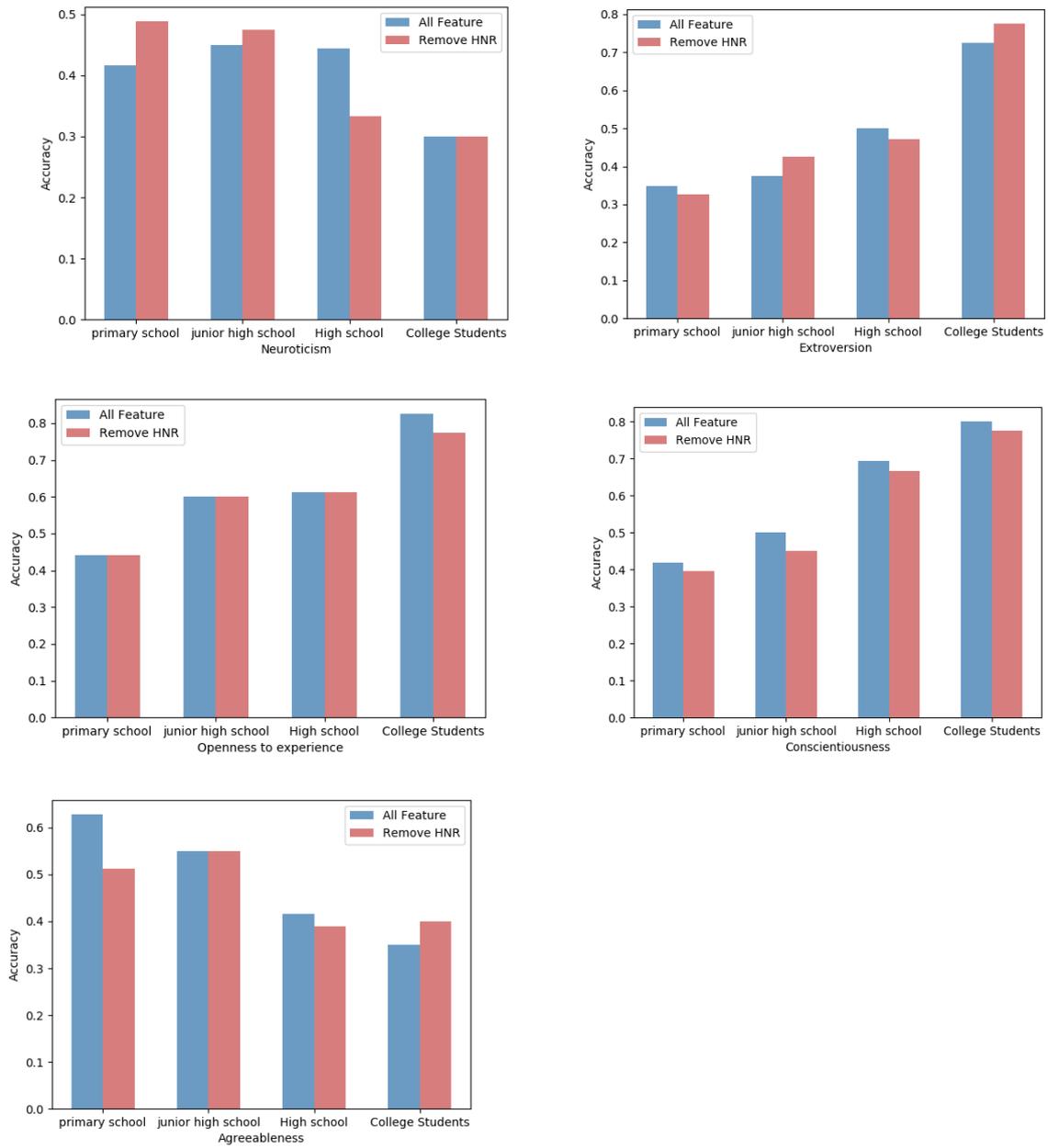


図 5.6.1 : 各年齢層で全特徴と HNR を取り除く結果の比較

属性	N	E	O	A	C
P 値	0.933	0.614	0.391	0.015	0.548

表 5.6.2: すべての特徴と HNR を取り除く場合での差の t 検定

確率が変わっていない。

3、外向性について、小学生と高校生の外向性の正確率が低下し、中学生と大学生の外向性の正確率が向上した。

4、開放性について、高校生以下の開放性の正確率が変わっていない。大学生の開放性の正確率が低下した。

5、勤勉性について、小学生と高校生の勤勉性の正確率が低下し、大学生の勤勉性の正確率が向上した。高校生の勤勉性の正確率が変わっていない。

第6章 考察

上記に音声特徴を 1 つずつ取り除いた再学習した結果を分析したところ、それぞれの音響特徴がパーソナリティ因子への貢献度が異なり、年齢での差異もあることが分かった。

詳しく分析した結果を以下に示す：

1、F0 はパーソナリティ分類の開放性と勤勉性に貢献がある。開放性への貢献は年齢層と特に相関がない。勤勉性への貢献は年齢が上がるとともに低下している。

2、MFCC はパーソナリティ分類の開放性と協調性の分類に貢献がある事がわかった。MFCC は成人期の開放性分類と全年齢層の協調性分類に貢献している。

3、RMS Energy は開放性と協調性の分類に貢献している。そのなか、特に成人期あるいは成人期に近い年齢層の人の開放性と高校生以下の協調性の分類に貢献している。

4、ZCR は開放性と協調性の分類に貢献している。年齢層に分けて分析すると、高校生以上の開放性分類と高校生以下の協調性に貢献している。また、年齢が高いほど、ZCR は開放性の分類への貢献が大きくなる傾向があり、協調性の

分類への貢献が小さくなる傾向がある。

5、HNR は開放性、協調性と勤勉性に貢献している。年齢層に分けて分析すると、HNR は成人期の開放性、全年齢層の協調性、高校生以下の勤勉性に貢献している。

異なる年齢層のパーソナリティ分類に対する音響特徴の貢献が異なるため、パーソナリティ分類モデルでは、異なる年齢層の被験者に対して、Big Five 人格の異なる項目を測定するときに、異なる音響特徴集を採用する価値がある。今までの研究では、研究者たちはあらゆる年齢の人々のパーソナリティ分類に同じ音響特徴集を使用しようとした。将来の研究では、年齢やパーソナリティ属性に応じて異なる音響特徴を利用することは、より良い結果を得られる可能性がある。

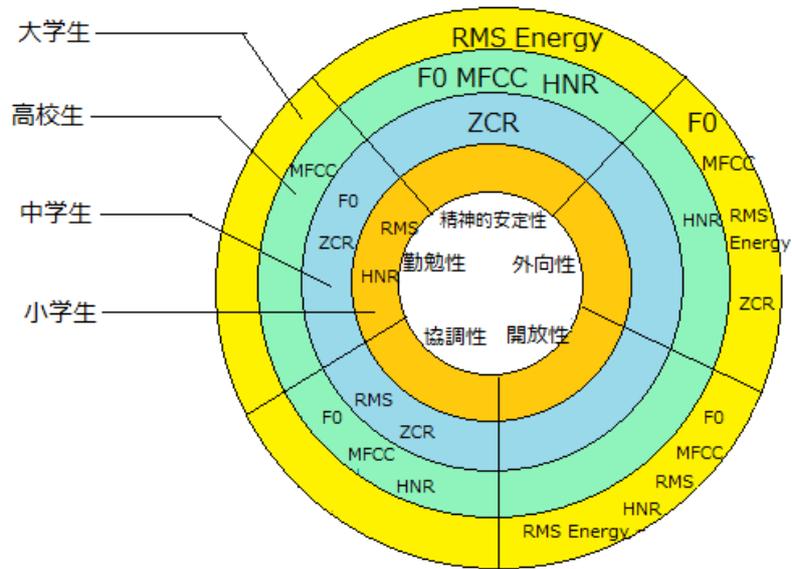


図 6.1 : 各年齢層でプラスの影響が大きい特徴の分布

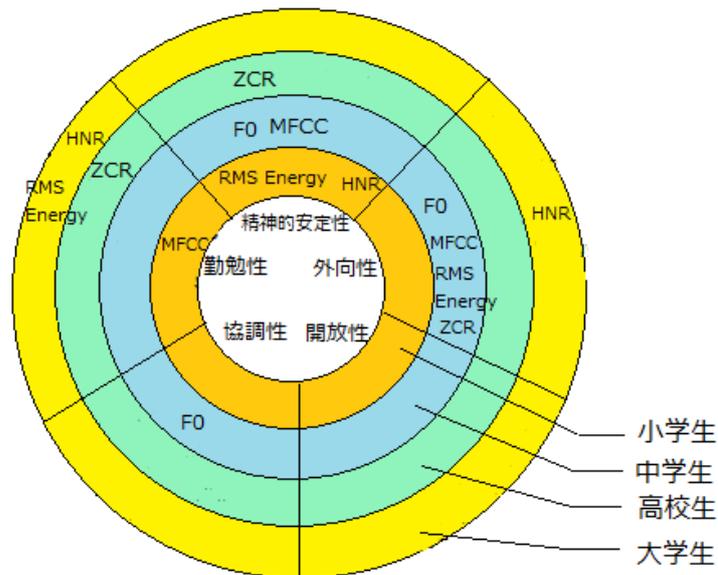


図 6.2 : 各年齢層でマイナスの影響が大きい特徴の分布

図 6.1 から見ると、大学生に対して、パーソナリティ分類への貢献が大きい特徴が一番多い。年齢が下がるとともに、貢献が大きい特徴が少なくなる。小学生に対して、貢献が大きい特徴が一番少ない。図 6.2 から見ると、中学生に対して、

マイナスの影響が大きい特徴が一番大きい。

これは、子供は変声期で音響特徴が変わっているからです。子供の変声期期間のばらつき、個人による変声のレベルの違いも原因になると考えている。これ以外、先行研究によって、青春기는パーソナリティ特性の変化の主要な期間であることがわかった。青春期中のパーソナリティの変化が成人期の変化より多いことも、子供のパーソナリティの分類効果が良くない原因になると考えている。

第7章 おわりに

本研究では、音響特徴とパーソナリティに関する因子および年齢との関連を分析した。その結果、音響特徴量によってパーソナリティ分類への貢献は異なり、その貢献度も年齢にも関連があることがわかった。したがって、将来では、異なる年齢と Big Five の 5 つのパーソナリティ属性に応じて、異なる音響特徴集を利用してパーソナリティを分類するモデルを研究する価値はある。本研究の異なる年齢でのパーソナリティ分類への音響特徴の貢献結果は将来の多特徴集を利用する分類モデルに知見を提供できると考えられている。

本研究は、年齢層が低いほど、パーソナリティ分類の正確率が低い傾向を示している。それは、少年期では、子供の発話・発声器官はまだ発達中で、音響特徴だけではなく、言語能力などに不確定な要素と、パーソナリティはまだ定着していない要素などがあると考えられる。

本研究について、改善できる点はデータベースの品質と大きさである。本研究で利用した録音のいくつはノイズが入ってる。できるだけ除去したが、残るノイズも実験結果に影響している。また、より良い深層学習モデルを訓練する場合、データ量の拡充は必要であると考えられる。

謝 辞

本論文をまとめるにあたりまして、指導教員の党建武教授にいろいろご指導とご助言をいただき、心よりお礼申し上げます。党研究室の皆様には、研究について色々なアドバイスを頂きまして、深く感謝いたします。特に、博士の Li Si Xia 先輩とすでに卒業した Zhai Yu 先輩からたくさんの指導を頂き、深く感謝いたします。また、研究の過程でプログラミングについて色々な問題がありまして、解決方法を教えていただいた吉田研の Chang Hen Yuan さんに深く感謝いたします。最後に、学生生活を手伝っていただいた友達と両親にご礼を申し上げます。

参考文献

- [1] Goldberg, L. R. An alternative ‘description of personality’: The Big-Five factor structure, *Journal of Personality and Social Psychology* 59, 1216–1229.(1990)
- [2] Komarraju, M., Karau, S. J., Schmeck, R. R. The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, vol. 51(4), pp.472-477. (2011)
- [3] 村上宣寛, 村上千恵子, 主要 5 因子性格検査の世代別標準化, *性格心理学研究* 2017, pp.1345-3629.(2017)
- [4] Goldberg LR, The structure of phenotypic personality traits, *American Psychologist*,vol.48 ,no.1,pp. 26–34.(1993)
- [5] Digman JM, Personality structure: Emergence of the five-factor model, *Annual Review of Psychology*, vol.41, pp.417–440. (1990)
- [6] 齊藤崇子, 性格特性用語を用いた Big Five 尺度の標準化, *九州大学心理学研究*, vol.2, 135-144.(2001)
- [7] 和田さゆり,性格特性用語を用いた Big Five 尺度の作成, *The Japanese Journal of Psychology*, vol.67, No.1, 61-67.(1996)
- [8] 内田 照久, 音声の韻律的特徴と話者のパーソナリティ印象の関係性, *音声研究*, vol.13, no.1, 17-28.(2009)
- [9] Donna Erickson. *Personality Judgments based on Speaker’s Social Affective Expressions*. Springer Nature Switzerland AG. (2018)
- [10] Gelareh Mohammadi. *Automatic personality perception: Prediction of*

- trait attribution based on prosodic features. *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 273-284.(2012)
- [11] Tim Polzehl, Florian Metze, Automatically assessing acoustic manifestations of personality in speech, 2010 IEEE Spoken Language Technology Workshop, pp.7-12.(2010)
- [12] Hu, Chao & Gen-Yue, FU. Perceive One's Character through His Voice: The Relationship between Speech Spectrum and Personality Traits. *Advances in Psychological Science*. vol.19, no.6, 809-813.(2011)
- [13] Roberts, B. W., Walton, K. E., & Viechtbauer, W. Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), pp.1-25.(2006)
- [14] 依田 明, 飯嶋 一恵, 出生順位と性格, 横浜国立大学教育紀要, vol.21, pp.117-127.(1981)
- [15] Wagner Jenny, Lüdtke Oliver & Robitzsch, Alexander. Does Personality Become More Stable With Age? Disentangling State and Trait Effects for the Big Five Across the Life Span Using Local Structural Equation Modeling. *Journal of Personality and Social Psychology*, vol.116. pp.666-680. (2019)
- [16] Sivrikova, Nadezhda & Artemyeva, Natalya & Sokolova, Nadezhda & Moiseyeva, Elena & Borodina, Vera & Terekhov, Anton. Features of the personality of the people identifying themselves with different generations. *SHS Web of Conferences*. vol.66(3):01003 (2019)
- [17] Alegre, Albert & Pérez-Escoda, Núria & Lopez, Elia. The Relationship Between Trait Emotional Intelligence and Personality. Is Trait EI Really Anchored Within the Big Five, Big Two and Big One Frameworks?. *Frontiers in Psychology*. 10. (2019)
- [18] Zou Rong, Tian Yuan, Stability and Change: Development of the Big Five Personality Domains in Western Children and Adolescents, *Journal of Psychological Science*, vol.39, no.4, pp.914-920.(2016)
- [19] 木下富雄, 社会心理学から見たパーソナリティ研究, パーソナリティ研究, vol.13, no.1, pp120-125.(2004)

- [20] 西尾正輝, 新美成二, 加齢に伴う話声位の変化, 音声言語医学, vol.46, 136-144.(2005)
- [21] 堀紗矢香, 田中章浩, 音声の基本周波数と抑揚が知覚年齢に及ぼす効果, 2014 年度日本認知科学会, vol.31 pp694-697.(2014)
- [22] 粕谷英樹, 森大毅, 加齢による声の音響的变化に関する通時的研究, 日本音響学会講演論文集
- [23] Zhai Yu, パーソナリティと音声コミュニケーション能力との関連についての研究, JAIST M-IS 2019 年度(R01).(2019)
- [24] Luxing Wang, A Survey of Chinese Language Testing in China in the Past Fifteen Years, *Modern Linguistics*, vol.7, no.01, pp.61-69.(2019)
- [25] Robert R. McCrae, Paul T. Costa Jr, Qin-Yu Zhu, A contemplated revision of the NEO Five-Factor Inventory, *Personality and Individual Differences*, vol.36, no.3, pp587-596.(2004)
- [26] Schuller, Björn , Steidl, Stefan, Batliner, Anton, The Interspeech 2009 Emotion Challenge. *Proc. Interspeech*. pp.312-315.(2009)
- [27] Guang-Bing Huang, Qin-Yu Zhu, Chee-K heong Siew. Extreme learning machine: Theory and applications. *Neurocomputing* 70. pp.489-501.(2006)
- [28] G. Huang, H. Zhou, X. Ding and R. Zhang, Extreme Learning Machine for Regression and Multiclass Classification, *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*. vol.42, no.2, pp.513-529. (2012)