

特別研究報告

題目

動画像視聴ユーザの脳波情報を用いた
機械学習による QoE 推定手法の実装と評価

指導教員

村田 正幸 教授

報告者

北尾 花純

平成 31 年 2 月 12 日

大阪大学 基礎工学部 情報科学科

動画像視聴ユーザの脳波情報を用いた
機械学習による QoE 推定手法の実装と評価

北尾 花純

内容梗概

近年、ネットワーク上を流れるトラフィックの増加は著しく、中でも動画像（ビデオ）コンテンツによるトラフィックは全 IP トラフィックの大半を占めている。ビデオコンテンツを提供するサービスやアプリケーションの設計においてはユーザのサービスに対する体感品質 (Quality of Experience; QoE) の向上は重要な課題である。近年では、動画像ストリーミングサービスにおいて QoE 情報を利用することにより、ユーザの QoE を向上するようなビットレート選択を行う手法の研究が行われるようになってきている。

ビットレートの選択にユーザの QoE を利用するためには、そのユーザ個人に適した QoE の測定が実時間でできることが必要となる。しかしながら、従来用いられている QoE の測定方法の多くは、通信品質のみに基づいてユーザの QoE を推定するもの、あるいはユーザにアンケートを取り、ユーザ自身が知覚した QoE を自己申告したデータを利用するというものであり、ユーザの個人差や実時間での測定という要件を満たすことは困難である。

我々はこれまでにパブリック EEG データセットを元にサポートベクターマシンと遺伝的アルゴリズムを組み合わせ、ユーザ自身の EEG データから感情推定を行う手法を提案してきた。本報告では動画視聴中のユーザの EEG データの収集実験を行い、上記手法を用いて QoE の推定を行なった結果を報告する。QoE は感情推定において分類対象となる喜怒哀楽といった強い感情に比べると分類すべき状態同士の差異が小さく、分類が困難となると考えられる。収集した EEG データを用いて分類モデルを学習した結果、3 クラスの QoE 分類における推定精度は 40.9%であった。

主な用語

QoE 推定、EEG (Electroencephalogram)、遺伝的アルゴリズム、サポートベクターマシン

目次

1	はじめに	5
2	関連研究	7
2.1	EEGによるユーザ体感品質(QoE)の評価	7
2.2	遺伝的アルゴリズム	7
2.3	サポートベクターマシン	8
3	ユーザの脳波情報を用いたQoEの推定手法	11
3.1	脳波の計測方法	11
3.2	プリプロセッシング	12
3.3	特徴量抽出	14
3.3.1	周波数帯域別パワー比	14
3.3.2	パワースペクトル密度	14
3.3.3	ウェーブレット変換	15
3.4	特徴量選択	16
3.5	QoE分類モデルの学習	17
3.6	QoEの推定	18
4	脳波情報を用いたQoE推定手法の性能評価	19
4.1	脳波の収集実験	19
4.2	QoE推定精度の評価結果と考察	21
5	終わりに	24
	謝辞	25
	参考文献	26

目 次

1	SVMによる2クラスパターン識別器の構成	9
2	多クラスSVMの構成法	10
3	EPOC+の外観	11
4	EPOC+のチャンネル配置	11
5	バターワースバンドパスフィルタのパラメータ図解	13
6	DWTを用いた特徴量の抽出	16
7	実験実施手順	21
8	QoEの平均推定精度	23
9	被験者ごとの推定精度	23

表 目 次

1	EEG の周波数成分分類	12
2	バターワースバンドパスフィルタのパラメータ	13
3	使用したバターワースバンドパスフィルタのパラメータ	14
4	ウェーブレット変換の各レベルでの周波数帯	15
5	サポートベクターマシンのパラメータ	18
6	ビデオクリップの品質	20
7	QoE の平均推定精度	23

1 はじめに

近年の動画像ストリーミングサービスの普及は著しく、ネットワーク上を流れるトラフィックの量は年々増加の一途を辿っている。世界での年間 IP トラフィック量は 2022 年までに 4.8 ZB に達し、内 82% を動画像 (ビデオ) コンテンツによるトラフィックが占めると予想されている [1]。

ビデオコンテンツを提供するサービスやアプリケーションの設計においてはユーザのサービスに対する体感品質 (Quality of Experience; QoE) の向上が重要な課題である。近年では、動画像ストリーミングサービスにおいて QoE 情報を利用することにより、ユーザの QoE を向上するようなビットレート選択を行う手法の研究が行われるようになってきている [2, 3]。

様々なモバイル端末が普及し、世界中のあらゆる場所からインターネットにアクセスすることができる現代において、ユーザのサービスの利用環境は多種多様である。それ故に、動画像視聴中のユーザの QoE は時々刻々と変動する可能性がある。また、ユーザごとに体感品質に影響を及ぼす要因は異なり、あるユーザにとっては体感品質を大きく変化させる要因が他のユーザにとっては体感品質にほとんど影響を与えないということも考えられる。従って、ビットレートの選択にユーザの QoE を利用するためには、そのユーザ個人に適した QoE の測定が実時間で行えることが必要となる。しかしながら、従来用いられている QoE の測定方法の多くは、通信品質のみに基づいてユーザの QoE を推定するもの、あるいはユーザにアンケートを取り、ユーザ自身が知覚した QoE を自己申告したデータを利用するというものであり、ユーザの個人差や実時間での測定という要件を満たすことは困難である。

近年では QoE と人の認知過程との関連を調査するための研究として、サービス体験中のユーザから直接計測できる情報である、表情、発汗、脳の血流変化、脳波、ユーザの具体的な操作 (動画再生中断) などからユーザの認知過程を反映した QoE の推定を行う研究が行われるようになってきている [4-7]。中でも脳波 (EEG; Electroencephalogram) と機械学習を用いた手法は高い精度を実現しているものがあり、QoE 推定に適していると考えられる。我々はこれまでにパブリック EEG データセットを元にサポートベクターマシンと遺伝的アルゴリズムを組み合わせ、感情推定を行う手法を提案してきた [8]。

本報告では動画視聴中のユーザの EEG データの収集実験を行い、[8] の手法を用いて QoE の推定を行なった結果を報告する。先行研究 [8] でのパブリックデータを利用した感情推定では分類のクラスを positive/negative の二値で実施したが、本報告ではビットレートの選択の判断材料として用いることを考慮し、Good、Normal、Bad の 3 クラスでの分類を行う。また、先行研究 [8] において分類器学習に用いる特徴量の個数や種類が分類精度に大きく寄与していることを踏まえ、本報告においても特徴量個数や種類による精度への影響を調査する。

本研究報告の構成は以下の通りである。2章では脳波を用いた QoE の推定についての関連研究について述べ、さらに本報告で実装する手法で用いる機械学習の手法について説明を述べる。3章では本報告で実装する手法について説明する。4章で脳波データの収集実験の内容および収集したデータを用いた推定精度の評価を行う。最後に、5章で結論と今後の課題について述べる。

2 関連研究

2.1 EEGによるユーザ体感品質 (QoE) の評価

近年、アプリケーションサービスの品質を表す指標としてユーザ体感品質 (QoE; Quality of Experience) が注目されている。一般的には QoE を測定する方法としてユーザ自身が選択回答するアンケートの形式が取られることが多いが、このアプローチでは得られる QoE の値がユーザの意識的な反応に依存しており、その根本にある潜在意識や認知のプロセスを無視している。そのため、近年では精神心理学に基づき、ユーザの無意識下の認知に焦点を当てた QoE 推定手法の研究が行われるようになってきている。精神心理学では人間の精神状態を身体的側面から観測するため、脳波や血流変化、心拍、眼球運動の解析が行われており、近年の QoE の推定方法の研究においても、これらの身体情報の解析法を応用しようとするものが多く見られる [4, 6, 7, 9]。中でも、脳波情報 (EEG; Electroencephalogram) は精神心理学的 QoE 推定の研究においてよく用いられており [6, 7, 9]、高い精度を実現している事例も存在していることから QoE の推定に適していると考えられる [7]。しかしながら、これらの既存研究においては QoE と生体情報間の関係性にユーザ毎の差異が存在する可能性を考慮していない。

我々はこれまでにパブリック EEG データセットを元にサポートベクターマシンと遺伝的アルゴリズムを組み合わせた機械学習アプローチにより感情推定を行う手法を提案し、その考察の中で被験者間でその推定精度に大きく差が現れることを示してきた [8]。本報告では、実際に動画視聴中のユーザから計測した EEG データを用いて、文献 [8] の手法により QoE の推定を行なった結果を示し、推定精度に与える要因について考察を行う。

2.2 遺伝的アルゴリズム

遺伝的アルゴリズム (GA) は組み合わせ最適化のためのメタヒューリスティックアルゴリズムの一つである。解の構成要素のランダムな組み合わせを複数作成し、それぞれの組み合わせにおける目的関数を評価してその優劣によって組み合わせの選択・淘汰と優れた組み合わせの交叉を繰り返すことで最適解の近似解を発見する。GA における解の構成要素を遺伝子、要素の組み合わせを個体、目的関数の評点を適応度と呼ぶ。GA の一般的な計算手順は以下の通りである。

1. 個体数 N 個の初期個体群を生成する
2. 個体群の各個体の適応度を計算する
3. 適応度に応じて次世代に残す個体を N 個選択する

4. 次世代個体群の中の個体の組に対して一定確率で交叉を行う
5. 次世代個体群の中の個体に対して一定確率で突然変異を行う
6. 交叉・突然変異が行われた個体の適応度を再計算
7. 3.以降の操作を適応度が最大に達するか規定の世代数に達するまで繰り返す

一般的に、上記の 3. において個体を選択する方式のことを選択関数と呼ぶ。また、4., 5. における交叉の方式、突然変異の方式をそれぞれ交叉関数、突然変異関数と呼ぶ。GA を組合せ最適化問題に適用する際には、遺伝子、個体、適応度、選択関数、交叉関数、突然変異関数と交叉・突然変異の発生確率を問題に応じて定義する必要がある。

本報告では、EEG データから抽出される数多くの特徴量の中から、QoE とより関連性が高い優れた特徴量の組み合わせを発見するために遺伝的アルゴリズムを使用する。

2.3 サポートベクターマシン

サポートベクターマシン (以降、SVM とも表記) は教師あり学習を用いて 2 クラスのパターン識別器を構成する手法の一つである。訓練データがラベル毎に線形分離可能であると仮定し、図 1 のように各クラスのデータ群の境界に位置するデータサンプル、すなわちサポートベクターからのマージンが最大となる超平面を求めることにより、データのクラス決定境界を得る。また、複数の SVM を組み合わせることにより、3 クラス以上のパターン識別器を構成することも可能である。SVM を組み合わせる方法は図 2 のように、1 クラス対 1 クラスの SVM をクラスの組み合わせの数だけ構成して合成する One-versus-One (OvO) 方式と 1 クラス対他の全クラスの SVM をクラスの数だけ構成して合成する One-versus-Rest (OvR) 方式の 2 種類がある。

線形分離可能でないデータ群に対しては、特徴ベクトルを非線形な写像を用いて高次元に変換し、その空間において線形分離を行う手法が用いられる。一般に、データ群を線形分離可能にするためにはデータのサンプル数と同程度の次元の空間への写像が必要となるため、SVM を構成するための目的関数の解を得るには膨大な計算量を要する。しかし、目的関数は各特徴の内積にのみ依存するため、内積を計算することができれば計算コストを下げることができる。そのため、式 (1) のように写像前の特徴ベクトルから写像後の特徴ベクトルの内積を計算することができるカーネル関数 K がしばしば用いられる。このように、カーネル関数を利用して高次元空間への写像を実現しつつ計算コストを抑える手法のことを一般にカーネルトリックと呼ぶ。ここで、 \mathbf{x}_i は元のデータの特徴ベクトル、 $\phi(\mathbf{x}_i)$ は写像後の特徴ベクトルを表す。

$$\phi(\mathbf{x}_1)^T \phi(\mathbf{x}_2) = K(\mathbf{x}_1, \mathbf{x}_2) \quad (1)$$

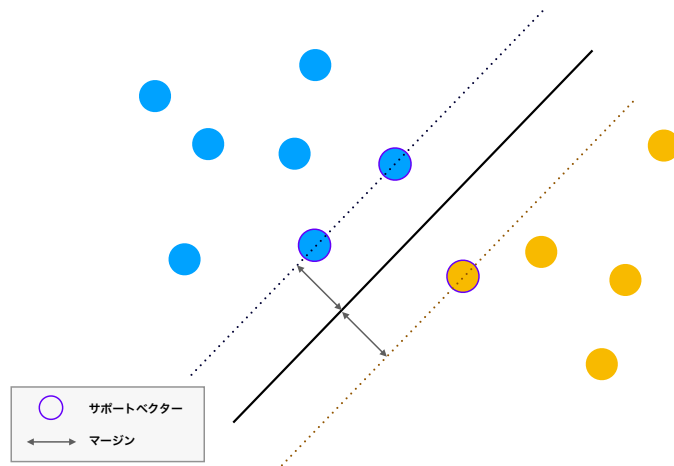


図 1: SVM による 2 クラスパターン識別器の構成

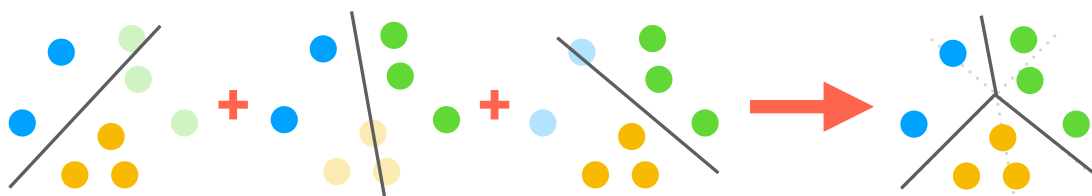
カーネル関数にはいくつか種類が存在するが、SVM においてよく用いられるものは多項式カーネル (式 (2))、RBF (Radial Basis Function) カーネル (式 (3)) などである。

$$K(\mathbf{x}_1, \mathbf{x}_2) = (1 + \mathbf{x}_1^T \mathbf{x}_2)^p \quad (2)$$

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp(-\gamma \|\mathbf{x}_1 - \mathbf{x}_2\|^2) \quad (3)$$

本報告では、式 (3) の RBF カーネルを用いた SVM により、QoE の 3 段階のクラスを分類する機械学習モデルを構成する。

One versus One(OvO)方式



One versus Rest(OvR)方式

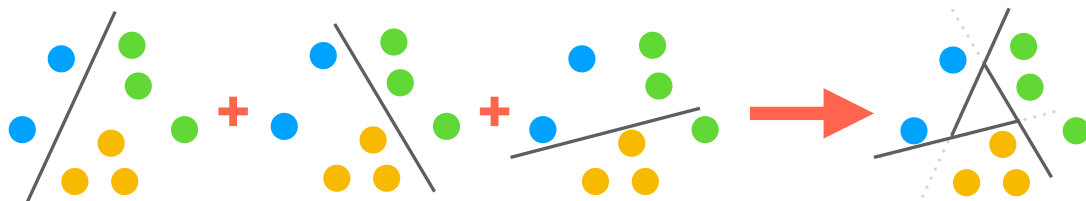


図 2: 多クラス SVM の構成法

3 ユーザの脳波情報を用いた QoE の推定手法

本報告における QoE 推定手法を説明する。具体的に、脳波情報の計測方法、解析方法、脳波情報から機械学習に用いる特徴量の抽出方法、分類モデルの学習方法について説明する。

3.1 脳波の計測方法

EEG センサーとして利用する EPOC+によって計測されるデータについて説明を行う。EPOC+は 14 チャンネルの電極とベース電位測定用の 2 チャンネル (図 4; EPOC+ユーザーマニュアル [10] より引用) を備えている。EEG データはデバイス内部で 2048sps (サンプリング/秒) でサンプリングされ、出力時に設定に応じて 128 または 256sps にダウンサンプリングされる。測定時の電圧解像度は 14bit (16bit で量子化された上で下位 2bit が削除される)、測定帯域幅は 0.16Hz-43Hz である。また、EPOC+はモーションセンサ (加速度計、ジャイロセンサ、磁力計) を備えており、設定に応じて 128Hz、64Hz、32Hz あるいはそれ以下のサンプリングレートで出力することができる。サンプリングしたデータは Bluetooth Smart により USB Bluetooth トランスミッターを介して PC に送信される。また、EPOC+から送信された EEG データおよびモーションデータを記録するソフトウェアとして同社 EmotivPro を使用する。

本報告では EmotivPro から CSV 出力した EEG データの内、チャンネル毎の 128Hz EEG Raw データ (原波形データ) を使用する。なお、実験で収集した EEG 情報は、計算機をオンライン状態にすると自動的に Emotiv 社のクラウドサービスにアップロードされるため、収集用の計算機は常時オフライン状態にして実験を行い、実験後にデータは別の記録装置に移動している。



図 3: EPOC+の外観

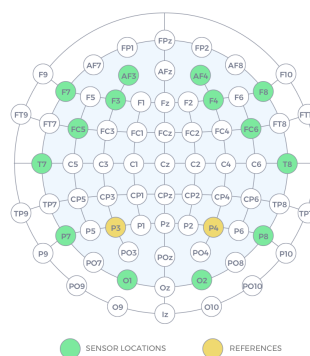


図 4: EPOC+のチャンネル配置

3.2 プリプロセッシング

計測した EEG データには脳活動と関係のない筋電位や脈拍、眼球運動などの生体現象によるアーチファクトや交流障害や電磁誘導など環境によるアーチファクト、測定電極の揺らぎによるドリフト電位等のノイズが混入している。そのため、特徴量の抽出を行う前にそれらの不要な信号を除去する必要がある。本報告では、EmotivPro のマニュアル [10] に記載の手法によりアーチファクトおよびノイズの除去を行う。この実装に置いては Matlab を使用した。EEG データから各チャンネルの信号の平均値を減算することにより各チャンネルに共通して混入しているノイズ成分を除去し、その後 0.16Hz 一次ハイパスフィルタにより低周波成分を減衰させ、ドリフト電位を除去する。

また、EEG の周波数成分は運動や認知による脳の活動で変化することが知られている。一般に用いられる周波数成分の分類を表 1 に示す。本報告においても周波数成分毎に特徴量の抽出を行う。各帯域毎の信号の抽出はバターワースバンドパスフィルタを使用する。バターワースバンドパスフィルタは特定帯域を通過させ、それ以外の帯域を遮断または減衰させる特性を持つフィルタである。バターワースフィルタは MatLab 上で実装する。実装に必要なパラメータを表 2 にまとめる。各パラメータは図 5 の通り、フィルターから通過させる周波数帯域と阻止領域での減衰量を指定するために用いる。パラメータの値は文献 [8] に記載のものを用いた。Astop1、Astop2、Apass は共通して 60dB、80dB、1dB に設定し、他のパラメータを各周波数帯域に合わせて変更して抽出を行なった。各周波数帯域毎のパラメータの値は表 3 の通りである。

表 1: EEG の周波数成分分類

名称	周波数帯
θ 波	4-8Hz
α 波	8-12.5Hz
β 波	12.5-30Hz

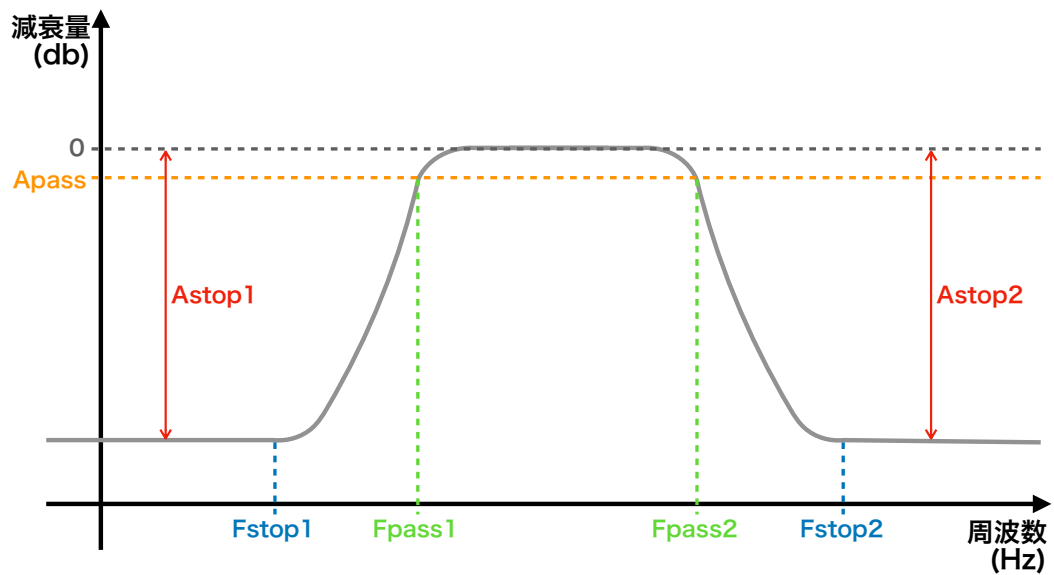


図 5: バターワースバンドパスフィルタのパラメータ図解

表 2: バターワースバンドパスフィルタのパラメータ

名称	概要
Astop1	1 番目の阻止帯域でのデシベル単位の減衰量
Astop2	2 番目の阻止帯域でのデシベル単位の減衰量
Apass	通過帯域で許容されるデシベル単位でのリップル量
Fstop1	1 番目の阻止帯域の始点周波数
Fstop2	2 番目の阻止帯域の始点周波数
Fpass1	通過帯域の始点周波数
Fpass2	通過帯域の終点周波数

表 3: 使用したバターワースバンドパスフィルタのパラメータ

周波数帯	Fstop1	Fpass1	Fpass2	Fstop2
全域	0.1	0.5	30	31
θ 波	3.8	4	8	8.2
α 波	7.8	8	12.5	12.7
β 波	12.3	12.5	30.5	31.7

3.3 特徴量抽出

収集した EEG データを時系列波形データとして機械学習アルゴリズムの入力とすることは計算量の観点から現実的でない。そこで、収集した EEG データから特徴量を抽出する操作が必要となる。本章では本報告で用いる特徴量の詳細および抽出方法について説明する。

各特徴量は EEG データを 2 秒の長さで分割したウィンドウ毎に計算し、それぞれに被験者の自己申告で得た QoE の値をラベルとして付与する。

3.3.1 周波数帯域別パワー比

バンドパワー (帯域パワー) は信号の平均パワーを表す指標である。本報告では各チャンネルの全域のバンドパワーに対する θ 波、 α 波、 β 波の割合を計算し、計 42 種類 (14 チャンネル \times 3 種類のバンドパワー) の特徴量を抽出する。

3.3.2 パワースペクトル密度

パワースペクトル密度 (以下、PSD) は単位周波数あたりのパワー分布を示した指標であり、EEG データを用いた感情推定及び QoE 推定において広く用いられている特徴量である [11-13]。信号から PSD を推定する手法はいくつか種類があるが、本報告では PSD の値を計算するための方法としてウェルチ法を用いる。ウェルチ法では、時系列の波形データを一定幅のセグメント (重ね合わせ可能) に分割し、セグメント毎に離散フーリエ変換を行なった後、パワーを計算し、最後にセグメント間で平均をとることにより PSD の推定を行う。本報告では各チャンネルの各周波数帯域の信号から計算された PSD から最大値、最小値、中央値、分散を計算して特徴量として用いる。よって、計 224 種類の特徴量が抽出される。

3.3.3 ウェーブレット変換

ウェーブレット変換はPSDと同じく周波数解析に用いられる手法であるが、時系列解析を同時に行うことができるという特徴を持つ。この手法はEEGデータの解析にも広く用いられている [14, 15]。ウェーブレット変換には連続ウェーブレット変換 (CWT; Continuous Wavelet Transform) と離散ウェーブレット変換 (DWT; Discrete Wavelet Transform) の2種類があるが、今回入力とするEEGデータは離散信号であるため、DWTを用いる。DWTは入力信号を半分にダウンサンプリングし、ウェーブレット関数に応じたバンドパスフィルタを適用するという操作を繰り返すことと等価である (図6)。この操作の適用回数はレベルと呼ばれ、各レベルで解析される周波数帯域は表4の通りである。

本報告では各チャンネルのDWTのレベル2から4の成分から最大値、最小値、平均値、分散を計算して特徴量として用いる。よって、計168種類の特徴量が抽出される。

表4: ウェーブレット変換の各レベルでの周波数帯

レベル	解析される周波数帯域	脳波成分区分
1	64-32Hz	ノイズ
2	32-16Hz	β 波
3	16-8Hz	α 波
4	8-4Hz	θ 波

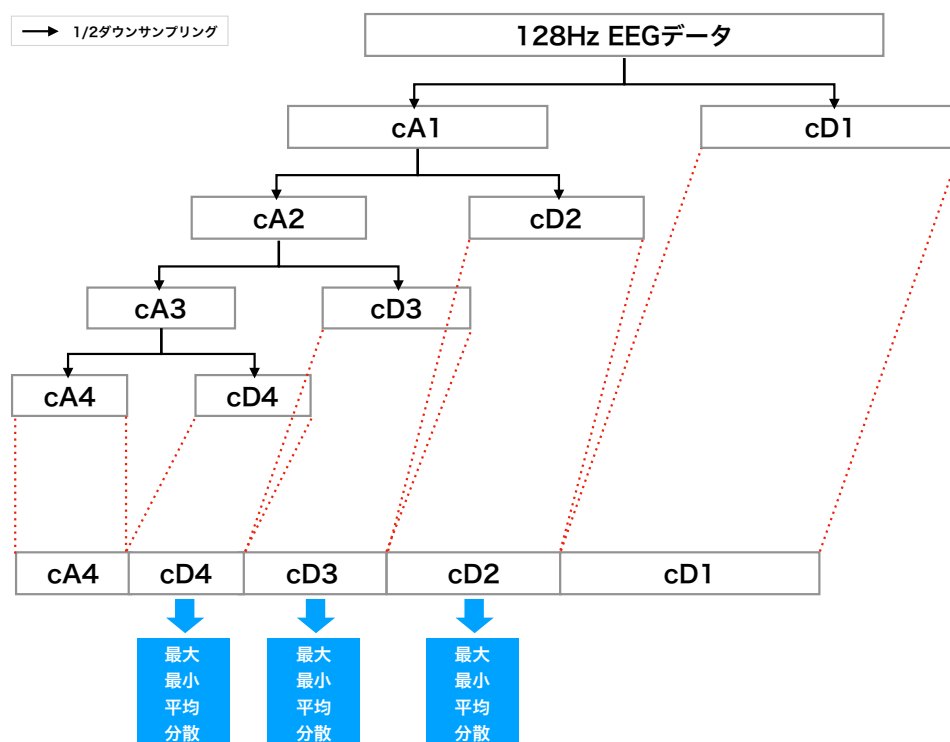


図 6: DWT を用いた特徴量の抽出

3.4 特徴量選択

SVM を用いた分類器の学習において、データサンプル数よりも多数の特徴量を入力した場合、過学習が発生し分類精度が低くなってしまふ恐れがある。そのため、抽出した特徴量 434 種類からモデルの学習に使用する特徴量を遺伝的アルゴリズム (Genetic Algorithm; GA) を用いて選択する。

GA の実装は Python 上で使用できる遺伝的アルゴリズムのライブラリである DEAP [16] を用いて行なった。選択する特徴量の個数は 3 種類、5 種類、10 種類でそれぞれ検証を行う。なお、GA の世代数は 1000、個体群の大きさは 100 としている。

GA の遺伝子は 434 種類の特徴量のいずれか 1 種類を示す 0 から 433 までの整数値、個体は 0 から 433 までの整数の順列である。選択する特徴量の個数に関わらず個体のサイズは 434 であり、適応度を計算する際に個体の先頭から指定の個数分の特徴量を取得して計算に用いる。それぞれの個体の適応度は RBF サポートベクターマシンによる 3 分割交差検証の精度とする。3 分割交差検証とは訓練データを 3 分割し、内 2 つを訓練データ、1 つをテス

トデータとして学習・評価を行うという操作をデータの分割を変更しながら3度行なって結果を平均するという検証手法である。適応度は個体の初期生成時に計算され、交叉または突然変異により構成遺伝子に変化があった場合のみ再計算される。

選択関数はトーナメント方式を用いる。トーナメント方式は個体群の中から一定数の個体をランダムに選び、その中で最も高い適応度を持つ個体を次世代に残す方式である。交叉関数は部分一致交叉法を用いる。部分一致交叉法は二つの個体から交叉を行う遺伝子をランダムに選択し、選択された遺伝子の並びが他方の個体と同じになるように個体内の遺伝子の並び替えを行うという交叉方法である。この方法では他の個体との遺伝子の交換は行われなため、個体内で遺伝子の重複が起こらず、選択する特徴量が重複することを防げる。突然変異関数は個体内の遺伝子配置をランダムに入れ替える転座法を使用する。この方法も交叉関数同様の理由で個体内の遺伝子の重複が発生しない。

交叉確率、突然変異確率はそれぞれ60%、1%に設定している。ただし、世代数が100の倍数のときのみ、突然変異確率を50%に引き上げる。これは、定期的に個体の編成を大きく変えることで個体群の多様性を維持し、局所的最適解への収束を防ぐためである。

3.5 QoE 分類モデルの学習

3.4章で選択した特徴量を入力として、被験者のQoEをgood, normal, badの3クラスに分類する機械学習モデルをRBF SVMを用いて学習する。

RBF SVMの実装はPython上で使用できる遺伝的アルゴリズムのライブラリであるscikit-learn [17]を使用して行なった。実装時に設定したパラメータを表5にまとめる。

モデルの学習は被験者毎に行う。データを動画像単位で2分割し半分を訓練データ、もう半分をテストデータとして訓練データのみを学習に用いる。テストデータはモデルの学習後、QoEの推定精度を評価するために用いる。なお、データの分割は、なるべく訓練データとテストデータに3種類全てのQoEラベルが含まれるように行う。データのラベル数の偏りにより、訓練データとテストデータの両方に3種類全てのラベルを振り分けることができない場合は訓練データに優先して振り分ける。

表 5: サポートベクターマシンのパラメータ

parameters	values	description
kernel	rbf	学習に用いるカーネル関数
C	1.0	誤分類の許容度
gamma	auto (1/特徴量数)	カーネル関数の係数
decision_function_shape	default(ovr)	多クラス分類時の SVM の構成方法
class_weight	balanced	クラス毎の C パラメータの重み

表 5 のパラメータのうち、class_weight はクラス毎のサンプル数の偏りによる影響を是正するためのものである。一般的に、分類すべき各クラスに対して訓練データのサンプル数に偏りがある場合、サンプル数の多いクラスからより大きなマージンが取られるため、決定境界がサンプル数の少ないクラスの側に押し寄せられる傾向がある。class_weight パラメータに balanced を指定した場合、クラス毎に全サンプルに占める当該クラスの割合の逆数が誤分類の許容度 C に乗算される。これにより、サンプル数の多いクラスの影響が低減し決定境界が適切な位置に設定されやすくなる。

3.6 QoE の推定

3.5 章で学習した分類モデルを用いて、未知のデータに対して QoE の推定を行う。

本報告では、3.1 章で使用しなかったテストデータを入力として QoE の分類を行い、分類結果と実際の QoE ラベルを比較した際の分類の正解率 (Accuracy) により QoE 推定手法の性能評価を行う。

4 脳波情報を用いた QoE 推定手法の性能評価

4.1 脳波の収集実験

Emotiv 社の EEG 計測用ハードウェア EPOC+および記録用ソフトウェア EmotivPro を用いて動画視聴中の被験者の EEG データを記録する。なお、人の生体情報を扱う実験のため、大阪大学の規定に従い事前に大阪大学研究倫理審査委員会へ申請を経て、研究機関の長より許可を受けて実験を実施している。

実験に使用する動画の準備

被験者の視聴するビデオクリップは文献 [8] で用意されたものを使用している。以下、用意されたビデオクリップの特徴、および準備の詳細について説明する。

- 長さが 4 分から 10 分
- 音声より映像の内容が中心となっている
- 内容が大衆向けである
- 高い品質のビデオクリップが利用可能である
- 品質の低下が QoE に大きく影響すると予想される

各被験者は 11 種類のビデオクリップの中からランダムに選択された 10 種類を視聴する。各ビデオクリップは元の品質からビットレート、フレームレートを変更した 5 段階の異なる品質のものを用意し、被験者が視聴する 10 種類のうち各品質が 2 つずつになるように準備する。なお、各品質の再生順はランダムである。品質の変更には ffmpeg を用いた。

ビデオクリップの 5 段階の品質の詳細を表 6 に示す。品質 1 から品質 5 にかけて段階的にビットレートあるいはフレームレートを落としたものとなっている。品質 1 はビットレートが 900kbps 以上、フレームレートが 24fps となっている。品質 2 ではフレームレートは品質 1 と同様に 24fps、ビットレートが 15 秒毎に 5 段階を経て 350kbps まで低下する。品質 3 ではビットレートは 900kbps 以上、フレームレートが 4fps で再生される。品質 4 ではフレームレートは 24fps、ビットレートが 15 秒毎に 5 段階を経て 250kbps まで低下する。品質 5 ではビットレートは 900kbps 以上、フレームレートが 2fps で再生される。ビットレートを段階的に変更しているのは前回の実験におけるビデオクリップ視聴の影響を低減するためである。

表 6: ビデオクリップの品質

品質	ビットレート (kbps)	フレームレート (fps)	再生数
1	< 900	24	2
2	~350	24	2
3	< 900	4	2
4	~250	24	2
5	< 900	2	2

被験者の募集

大阪大学学内情報システム KOAN にて大阪大学に所属する全学生を対象に被験者募集の掲示を行うことにより被験者を謝礼付きで募集した。掲示内容は以下の4点である。募集者の選定は先着順とし、合計25人の被験者が集まった時点で募集を終了した。応募辞退を含めて、合計21人の被験者を対象に実験を行った。

内容 いくつかの動画像を視聴している間の EEG データの計測

対象 大阪大学に属する学生

所要時間 1時間半から2時間

謝礼 3000円(交通費込)

実験の実施手順

実験の手順を図7に示す。実際に実験を行う際は、図7の手順に加えて動画視聴の合間に EPOC+の接触を調整する、被験者の体調を鑑みて数分の休憩を挟むなどの状況に合わせた調整を行なっている場合がある。

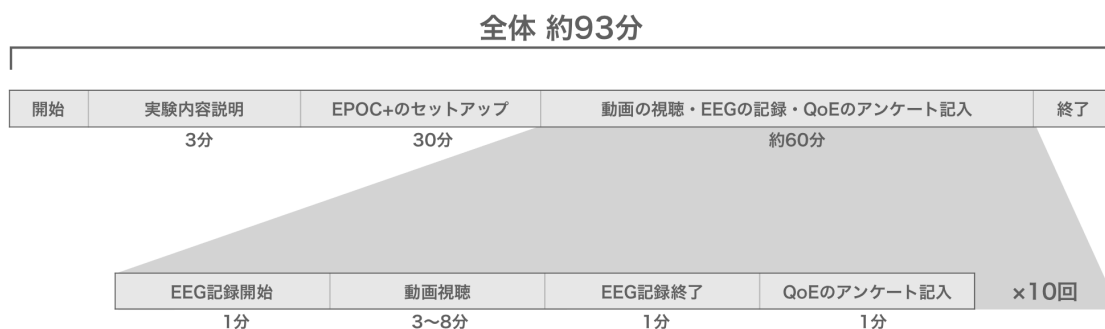


図 7: 実験実施手順

脳波の計測

脳波の計測はEPOC+と記録用のソフトウェア EmotivPro を用いて行う。EEG・モーションデータのサンプリングレートは双方とも 128sps (sample per second) である。データの記録は動画再生と同時に開始し、動画の終了後停止する。データを使用する際は、分類精度を向上させるために被験者の QoE が大きく変化していると考えられる部分、すなわち動画像の品質が大きく変化する 15 秒から 1 分 30 秒のデータのみを使用する。

学習ラベル情報の収集

QoE の分類モデル学習の際の目的関数 (ラベル) となる QoE の値を収集する。各動画の視聴後、被験者に対して動画の印象を good、normal、bad の三段階で評価してもらい、これを学習ラベルとして用いる。

4.2 QoE 推定精度の評価結果と考察

GA により選択された特徴量とランダムに選択された特徴量でそれぞれ RBF サポートベクターマシンを用いて QoE 分類モデルを学習し、精度を評価した。特徴量の選択手法と特徴量数ごとの平均分類精度を図 8 と表 7 に示す。ここで、分類精度とはテストデータが正しく推定される割合を示している。モデルの学習および評価は被験者毎に行なっているため、学習ラベル情報収集において QoE のラベルが 3 種類以下になった被験者のデータは使用していない。そのため、実際に使用したデータは全被験者 21 名中 19 名分となっており、それぞれの被験者についての分類精度および全被験者についての平均精度を調査した。

先行研究 [8] において、パブリックデータでの感情推定を行なった際の平均精度は 74.15% であったのに対し、今回の実験での平均精度は 40.7 % 前後に留まっている。これは、前研究における SVM の学習ラベルが positive と negative の 2 クラスであったことに対して、今回は 3 クラスの QoE ラベルを用いていることで分類問題の難度が高くなったこと、感情推定での分類対象である喜怒哀楽のような強い感情に比べて、満足かどうかという QoE の分類では、クラス間の特徴量の差異が小さく分類が困難であることが原因として考えられる。加えて、感情推定で使用したパブリックデータセットの収集環境と比較すると、今回の実験環境は、計測装置の精度の点、データ数の点、装置装着時の違和感や疲労など動画視聴時の QoE の変化以外の要因のコントロールが行えていなかった点で万全ではなかったと考えられる。これらを改善するためには、学習データの収集における環境について被験者数を増やす、使用する計測装置の精度を改善する、各動画視聴の間の休憩時間を長く設けて疲労による脳波への影響を低減するといった対策が考えられる。

特徴量選択手法として GA を用いた場合とランダムに選択した場合の比較では、GA を用いた結果が平均分類精度で 4.9% 上回った。この結果は、EEG での QoE 推定におけるチャンネルや特徴量の種類の選択が精度に与える影響が大きいことを示唆している。一方で、今回の実験においては特徴量の個数の違いによる精度差は 0.1% 以下であり、大きな差異は認められなかった。この理由としては特徴量の個数が分類精度に与える影響がそもそも小さい可能性が考えられるが、一方で先行研究での感情推定における結果では差異が表れていたため、特徴量個数の調査範囲が狭く、影響が観測できなかつた可能性も考慮する必要がある。

ここで、被験者毎の分類精度を図 9 に示す。特徴量の分類方法、使用する特徴量の個数により多少のばらつきはあるが、被験者間で最大 36.5% の精度差が存在する。この結果は、EEG による QoE の推定精度において個人差が比較的大きいことを意味している。また、特徴量の個数や種類の影響についても、個人差の存在が確認できた。すなわち、特徴量個数および種類の選択を行い、精度を向上する際には、個人ごとにチューニングを行うことが重要であると言える。

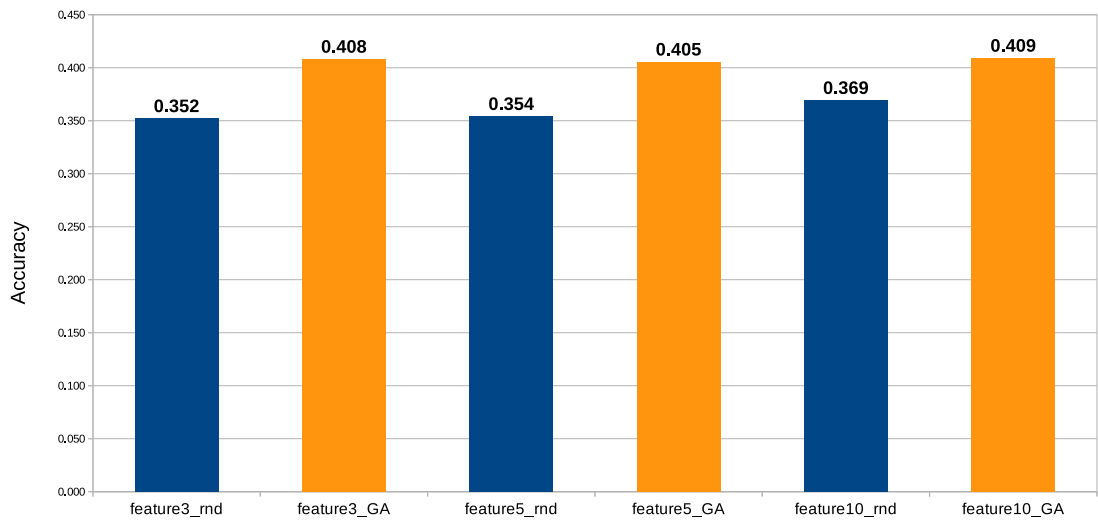


図 8: QoE の平均推定精度

表 7: QoE の平均推定精度

特徴量数_特徴量選択法	平均精度	分散
feature3 random	0.352	0.003
feature3 GA	0.408	0.007
feature5 random	0.354	0.002
feature5 GA	0.405	0.005
feature10 random	0.369	0.003
feature10 GA	0.409	0.008

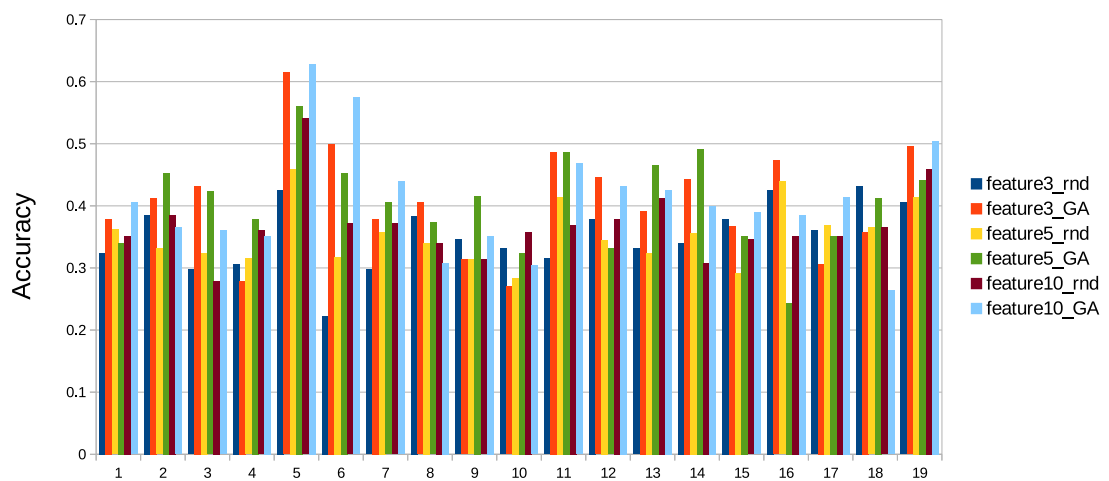


図 9: 被験者ごとの推定精度

5 終わりに

ユーザ体感品質 (QoE) の推定は動画像アプリケーションサービスプロバイダにとって重要な課題である。本報告では、我々がこれまでに取り組んできたユーザの EEG データを用いた感情推定手法を QoE 推定に適用し、被験者を対象とした実験を行った結果を報告した。

動画像視聴中の被験者の EEG データと動画像視聴後の QoE の評価回答を基に、GA と SVM を組み合わせた学習手法により QoE 分類モデルの学習と評価を行なった。学習に用いる特徴量の個数を変えながら、使用する特徴量を GA を用いて決定した結果、最も良い特徴量数を選んだ場合に、推定精度は平均 40.7% となった。一方で、ランダムに特徴量を選択した場合の推定精度は 35.8% であり、適切な特徴量を選択することが重要であることを示した。特徴量の個数に関しては、感情推定に用いたときは精度に大きく影響を与えていたが、今回の実験ではほとんど影響を与えないという結果となった。

今後の課題として、QoE の推定精度の向上を第一とし、そのために脳波情報の収集実験の改善を行う。本報告では一つの動画について一度の QoE 評価を行なっていたが、これはユーザの QoE 決定因子が時間経過により変動する可能性を考慮できていない。また、実際の動画像アプリケーションサービスでは、ユーザ毎にデバイスや環境が変化することにも留意する必要がある。今後 QoE 推定を実時間で行う手法についても検討を行う。

謝辞

本報告を終えるにあたり、日頃よりの確にご指導、ご教授くださいました大阪大学大学院情報科学研究科の村田正幸教授に深謝いたします。ならびに、本報告の作成に際し、終始懇ろにご指導頂きました大阪大学大学院経済学研究科の小南大智助教に心より厚く御礼申し上げます。また、平素よりご指導くださいました大阪大学大学院情報科学研究科の荒川伸一准教授、先導的学際研究機構の大下裕一准教授に深く感謝いたします。本報告の基本的なアイデアおよび機器やソフトウェアの扱いについて丁寧にご指導くださいました Jens Oetjen 氏に心より感謝申し上げます。最後に、脳波情報収集実験に協力して頂いた他、日頃より親身にご指導くださいました岩本真尚氏、細川侑嗣氏、松田拓己氏、実験の準備と機器の試用に協力してくれた杉浦満美氏をはじめとする村田研究室の皆様、実験の被験者をお引き受けくださった大阪大学の学生の方々、および家族、友人の皆様に感謝の意を表し、謝辞とさせていただきます。

参考文献

- [1] Cisco, “Cisco Visual Networking Index: Forecast and Trends, 2017-2022,” 3 February 2019. https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-741490.html#_Toc529314172.
- [2] D. Suh, I. Jang, and S. Pack, “Qoe-enhanced adaptation algorithm over dash for multimedia streaming,” in *The International Conference on Information Networking 2014 (ICOIN2014)*, pp. 497–501, IEEE, February 2014.
- [3] Masayoshi Iwamoto, “A rate control method for QoE improvement in video streaming services based on a human cognitive model,” Master’s thesis, Graduate School of Information Science and Technology, Osaka University, Feb 2019.
- [4] K. Ota, T. Hiraguri, and H. Yoshino, “A study on QoE estimation from heart rate variability using machine learning,” in *Proceedings of 2018 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, pp. 1–2, IEEE, May 2018.
- [5] U. Engelke, D. P. Darcy, G. H. Mulliken, S. Bosse, M. G. Martini, S. Arndt, J.-N. Antons, K. Y. Chan, N. Ramzan, and K. Brunnström, “Psychophysiology-based QoE assessment: A survey,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, pp. 6–21, February 2017.
- [6] U. Reiter and K. De Moor, “Content categorization based on implicit and explicit user feedback: combining self-reports with EEG emotional state analysis,” in *Proceedings of the Fourth international Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 266–271, IEEE, July 2012.
- [7] P. Arnau-Gonzalez, T. Althobaiti, S. Katsigiannis, and N. Ramzan, “Perceptual video quality evaluation by means of physiological signals,” in *Proceedings of the Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1–6, IEEE, June 2017.
- [8] Jens Oetjen, “Predict Video Streaming QoE with Brainwave Data Based Support Vector Machines Boosted by a Genetic Algorithm,” Master’s thesis, Graduate School of Information Science and Technology, Osaka University, August 2018.

- [9] J. Donley, C. Ritz, and M. Shujau, “Analysing the quality of experience of multisensory media from measurements of physiological responses,” in *Quality of Multimedia Experience (QoMEX), 2014 Sixth International Workshop on*, pp. 286–291, IEEE, December 2014.
- [10] EMOTIV, “EPOC+ user manual, 2018.” <https://emotiv.gitbook.io/epoc-user-manual/>.
- [11] S. Akramullah, “Video quality metrics,” in *Digital Video Concepts, Methods, and Metrics*, pp. 101–160, Springer, October 2014.
- [12] W.-L. Zheng and B.-L. Lu, “Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks,” *IEEE Transactions on Autonomous Mental Development*, vol. 7, pp. 162–175, May 2015.
- [13] M. Soleymani, M. Pantic, and T. Pun, “Multimodal emotion recognition in response to videos,” *IEEE Transactions on Affective Computing*, vol. 3, pp. 211–223, April–May 2012.
- [14] M. Murugappan, N. Ramachandran, and Y. Sazali, “Classification of human emotion from EEG using discrete wavelet transform,” *Journal of Biomedical Science and Engineering*, vol. 3, pp. 390–396, April 2010.
- [15] H. Adeli, Z. Zhou, and N. Dadmehr, “Analysis of EEG records in an epileptic patient using wavelet transform,” *Journal of Neuroscience Methods*, vol. 123, pp. 69–87, February 2003.
- [16] F.-A. Fortin, F.-M. De Rainville, M.-A. Gardner, M. Parizeau, and C. Gagné, “DEAP: Evolutionary algorithms made easy,” *Journal of Machine Learning Research*, vol. 13, pp. 2171–2175, July 2012.
- [17] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.