

# 基本感情の複合分析に基づく表情画像からの困惑深度推定 : Transformer を用いた特徴抽出と評価

石川 昂樹<sup>†</sup> 一色 夢香<sup>†</sup> 寺田 憲司<sup>†</sup> 遠藤 雅樹<sup>†</sup> 大野 成義<sup>†</sup>

<sup>†</sup> 職業能力開発総合大学校 電子情報専攻 情報通信ネットワークユニット 〒187-0035 東京都小平市小川西町 2-32-1

E-mail: † {b22302, b22305, k-terada, endou, ohno}@uitec.ac.jp

**あらまし** 急速に常態化した非対面の学習・作業環境での困惑状態の把握は作業停滞の早期発見や心理的安全性の確保を実現しうるデジタル社会に求められる技術である。そこで本研究では、困惑が「怒り」、「悲しみ」等の複数の基本感情の要素を含んでいる可能性に着目する。具体的には、作業時に取得した表情画像から学習済み深層学習を用いて複数の基本感情を抽出し、教師付き機械学習によって困惑を推定する手法を提案する。実験では7名の被験者に対しプログラミング課題を課し、Webカメラで表情を撮影した。表情画像からTransformerを用いて7つの基本感情スコアのデータセットを構築した。さらに、データセットには特徴量増加を施し、説明変数を生成した。プログラミング課題中に4段階の困惑深度のアンケートを実施し、目的変数とした。機械学習による困惑推定モデルを構築した結果、各困惑深度の平均F値は0.8882となった。研究から困惑は基本感情を組み合わせた複合感情により推定でき、被験者毎に困惑推定が可能であるという知見を得た。

**キーワード** パターン認識, 深層学習, 機械学習, 感情分類, 生体センシング

## 1. はじめに

近年、社会全体においてオンライン環境での作業が普及している。テレワークの導入率は2018年から2024年にかけて20%以上増加している[1]。また、教育現場においても大学生のオンライン授業の受講率は約80%に達している[2]。しかし、このようなオンライン環境での作業の普及に伴い課題も顕在化している。オンライン授業中に教員が受講生の反応が見えないという回答が約半数を占めており[3]、教員の約60%が学生の困惑が把握できないことが報告されている[4]。

特にオンライン環境におけるプログラミング初学者を対象とした授業実践では、学習者がつまずきの原因を言語化できず、質問機能を用いても適切な支援を求められないケースがあることが指摘されており[5]、オンライン環境下における潜在的な困惑の検知は困難を極める。そのため、非対面の学習作業環境での困惑状態の把握は、作業停滞の早期発見や心理的安全性の確保を実現しうるデジタル社会に求められる技術である。

しかし、困惑の表情の検出には課題がある。第一に、困惑の表情は他の思考中の表情と類似していることが示唆されており、その識別の複雑さが浮き彫りになっている[6]。第二に、学習者が困難に直面した際に生じる「困惑」が重要であるが、表情のみでは中立的な表情や他の感情と誤認識されやすい[7]。第三に、学習中の困惑は頻繁に発生するが、表情は微妙で検出しにくいことも報告されている[8]。

そこで本研究では、プログラミング作業時の表情から学習済み深層学習を用いて複数の基本感情を抽出し、

教師付き機械学習によって困惑深度を推定することを目的とする。

## 2. 関連研究

### 2.1. 困惑の表情特徴と既存検出手法の課題

Yasserら[9]の研究によれば、眉の引き下げと瞼の緊張の組み合わせは、困惑状態において有意に頻出し、口元の動きは、困惑が生じている際の「何か言いたげだが言えない」状態や、自己嘲笑的な反応を反映していると分類している。表情から困惑を推定できるとしており、これらの特徴はポール・エクマンが提唱する「怒り(angry)」、「嫌悪(disgust)」、「恐怖(fear)」、「驚き(surprise)」、「幸福(happy)」、「悲しみ(sad)」、「中立(neutral)」の7つの基本感情の特徴にみられるものと類似している。

一方で、既存手法の侵襲性とコストの問題、従来の混乱検出技術の多くは、脳波(EEG)や筋電図(EMG)といった生理学的センサを身体に装着する必要がある。これらは侵襲的であり、高価で、かつユーザーの動きを制限するため、日常的なEラーニング環境などでの利用には適していないと課題に挙げている。

### 2.2. 困惑検出における時間的変化の重要性

Pachmanら[10]は、困惑を単なる一瞬の表情としてではなく、時間的な広がりを持つ動的なプロセスとして捉えている。学習のきっかけとなる「建設的な混乱(Productive Confusion)」と、解決できずにフラストレーションへと移行する「非建設的な混乱(Unproductive Confusion)」に分類するとしている。困惑の検出には

Hesitation と呼ばれる意思決定や発話の前の「間」として現れる行動的シグナルを検出するためには、個別の特定の顔面微細表情の検出が必要であるとされている。

この関連研究から得られることは、一瞬の画像から得られる単なる分類モデルでは困惑推定が難しいということである。

### 2.3. 困惑状態の細分化の必要性

Hussain ら[11]の研究によれば、感情的干渉がある状況下では生理信号よりも顔表情の方が認知的負荷の困惑検出においてロバストであることが示されている。これは、生理反応が情動に強く反応してしまう一方で特定の顔面筋は認知的努力をより純粋に反映することを明らかにしている。但し、顔、生理信号などのマルチモーダル情報が増えると複数の情報源を統合することで一般的には検出精度が向上するが、本研究では感情的覚醒の強度が高まると統合してもなお検出精度が低下することが明らかになった。つまり、強い感情的干渉下では、単純なマルチモーダル統合だけではロバスト性を維持できない点が課題として浮き彫りになった。

この研究から、困惑の深度が高まった場合、どのような感情の変化がみられるかを明らかにすることが困惑推定で必要となることを特定した。

## 3. 提案手法

### 3.1. 概要

本研究では、困惑が基本感情の複合感情である可能性に着目し、プログラミング作業中の作業者の表情画像から困惑を推定する手法を提案する。

従来の研究では、表情画像から困惑推定モデルを構築する際、「困惑」状態をラベル付けした学習画像データが必要である。しかし、十分な困惑画像データの収集は大きな労力が必要であるという課題がある。ここで、7つの基本感情分類には学習済み深層学習モデルが存在することに注目する。本研究では、画像から直接「困惑」を学習させるのではなく、既に確立された学習済み深層学習モデルから得られるスコア結果から困惑を分類する手法を提案する。

データ取得から評価までの提案手法を示す(図1)。Webカメラを用い、被験者の表情を撮影し、表情画像を取得する。表情画像から深層学習を用い、7つの基本感情スコアを取得する。このスコアからスライドウィンドウを用い、統計的特徴量増加を行った。これを説明変数とした。プログラミング作業中に4段階の困惑深度アンケートを実施し、これを目的変数とした。説明変数と目的変数をタイムスタンプで結合し、特徴量を生成した。教師付き機械学習を用い、困惑推定モ

デルを構築する。モデルは評価および複合感情分析を行う。

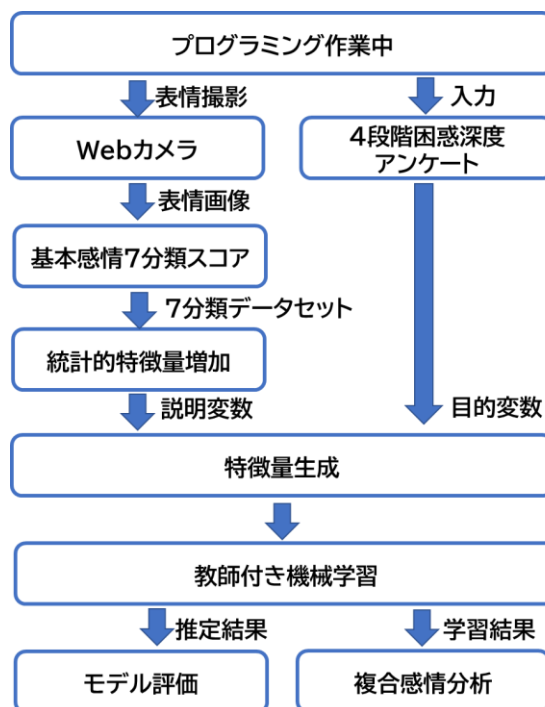


図1 提案手法のブロック図

### 3.2. 特徴量

「2.2. 困惑検出における時間的変化の重要性」より、困惑は一瞬の表情としてではなく、時間的なプロセスとして捉える必要がある。そこで、機械学習を適用する際は統計的特徴量増加によって時間的変化を特徴量の中に持たせる。

基本感情の複合感情から困惑推定を行うために、表情画像から7つの基本感情スコアを取得する(式1)。

$$E_{x,t} = D_t(I)x \quad (\text{式 1})$$

$E_{x,t}$ : 基本感情スコア (0.0~1.0)

$t$ : 時間

$D_t$ : 深層学習

$I$ : 表情画像

$x$ : an(angry), di(disgust), fe(fear), su(surprise),  
ha(happy), sa(sad), nu(neutral)

ある時間における7つの基本感情スコアを示す(式2)。

$$E_{B,t} = \{E_{an,t}, E_{di,t}, E_{fe,t}, E_{su,t}, E_{ha,t}, E_{sa,t}, E_{ne,t}\} (\text{式 2})$$

$E_{B,t}$ : ある時刻における7つの基本感情スコアスライドウィンドウを用いた(式3)。

$$E_S = \{E_{B,t}, E_{B,t-1}, \dots, E_{B,t-w}\} (\text{式 3})$$

$E_S$ : ローデータ

$w$  : ウィンドウサイズ

取得した $E_S$ に対し、最小値、最大値、中央値、平均値、分散、標準偏差、尖度、歪度の統計的特徴量増加を行った(式4)。

$$F_g = \{F_{min}, F_{max}, F_{med}, F_{mea}, F_{var}, F_{std}, F_{kur}, F_{ske}\} \quad (式 4)$$

$F_g$  : 特徴量

$F_{min}$  :  $E_S$ の最小値

$F_{max}$  :  $E_S$ の最大値

$F_{med}$  :  $E_S$ の中央値

$F_{mea}$  :  $E_S$ の平均値

$F_{var}$  :  $E_S$ の分散

$F_{std}$  :  $E_S$ の標準偏差

$F_{kur}$  :  $E_S$ の尖度

$F_{ske}$  :  $E_S$ の歪度

ローデータと統計的特徴量増加によって説明変数に適用する特徴量数は63となった。機械学習を用い、予測結果を評価する(式5)。

$$P_e = M_{RF}(E_{B,t}, F_g) \quad (式 5)$$

$P_e$  : 困惑深度

$M_{RF}$  : 機械学習

### 3.3. 評価方法

#### 3.3.1. 機械学習モデルの評価

機械学習モデルの評価指標はF値(F-score)を適用する。F値は混同行列から求めることができる「適合率(Precision)」と「再現率(Recall)」のバランスを評価するための重要な指標である。特にデータセットのクラスに偏りがある場合や、誤検出と見逃しの両方を考慮したい場合に利用する。混同行列とは機械学習の分類問題において、モデルの予測結果と実際の正解値を対照させた表のことである。モデルが「どのクラスをどの程度正しく予測できたか」、あるいは「どのクラスと間違えやすいのか」を評価するための指標である。例えば二値分類(Positive / Negative)の場合、行に実際の正解クラス、列に予測クラスを配置する(表1)。

表1 混同行列

		予測	
		Positive	Negative
実際	Positive	TP	FN
	Negative	FP	TN

混同行列を構成する4つの要素は以下の通りである。

- TP (True Positive) : 実際も正で、予測も正
- TN (True Negative) : 実際も負で、予測も負

- FP (False Positive) : 実際は負だが、予測は正
  - FN (False Negative) : 実際は正だが、予測は負
- これら4つの要素を利用した適合率はモデルが「正(Positive)」と予測したもののうち、実際に正であったものの割合である。すなわち、「予測の正確さ」に焦点を当てた指標である(式6)。

$$Precision = \frac{TP}{TP + FP} \quad (式 6)$$

再現率は、実際に「正」であるもののうち、モデルが正しく正と予測できたものの割合である。すなわち、「見逃しの少なさ」に焦点を当てた指標である(式7)。

$$Recall = \frac{TP}{TP + FN} \quad (式 7)$$

F値とは適合率と再現率により求めることができる。適合率と再現率はトレードオフの関係にある。これら2つの指標を統合し、一つの値でモデルの性能を評価する指標がF値である(式8)。

$$F\text{-score} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (式 8)$$

#### 3.3.2. 複合感情分析の評価

複合感情分析の評価には、Explainable AI(XAI)手法の1つであるSHAP[12]を適用する。SHAPはゲーム理論のシャープレイ値に基づき、各特徴量が予測値の偏差に与える寄与度を算出する手法である。モデル全体において、「どのデータが分類に重要であるか」また、「互いにどう影響しているか」を分析するものである。これにより、ブラックボックス化しやすい機械学習モデルにおいて、個々の予測に対する特徴量の影響度を定量的に評価することが可能になる。本研究ではViolin Summary Plot(図2)を適用する。左側の縦軸は特徴量を示し、寄与度が大きい順に上から整列している。横軸はSHAP値を示し、0より右側は分類値に対する正の影響を示し、左側は負の影響を示す。各点は個々のサンプルを表し、色は特徴量の値(赤:高い、青:低い)を表している。つまり、赤い点が右側に多いほど予測に有効な特徴量となる。図2では「嫌悪」、「怒り」、「恐怖」が予測に有効な特徴量となる。

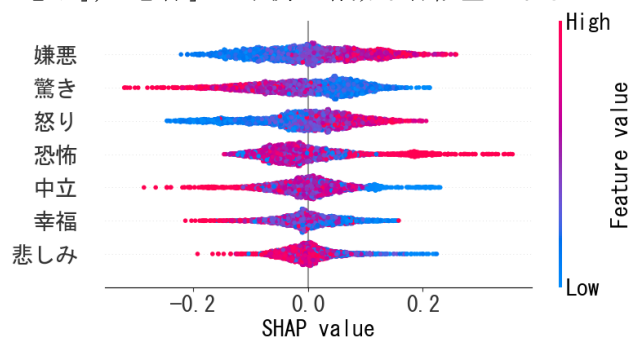


図2 Violin Summary Plot

## 4. 実験

### 4.1. データ収集方法

プログラミング作業中の作業者の表情画像を取得することを目的とし、プログラミング課題を課した。18歳から22歳の大学生7名が参加した。プログラミング課題に対し、内容は講義として履修したことがあるものに限定した。

本研究は、職業能力開発総合大学校倫理審査委員会の承認を受けた。

### 4.2. 実験環境

本実験環境を図3に示す。実験は、PCと2台のディスプレイ（サイズ：24インチ、解像度：1920×1080）を使用した。2台のディスプレイは横に並び、各ディスプレイの画面を左右2分割し、計4つの領域でウィンドウを常時表示させた。ウィンドウはそれぞれ左から課題内容を表示する課題提示画面、解答コードを入力するコードエディタ、教科書などの参照資料、困惑深度アンケート画面である。コードエディタ画面の上部に記録解像度1080PのWebカメラ（Angetube社967）を設置した。

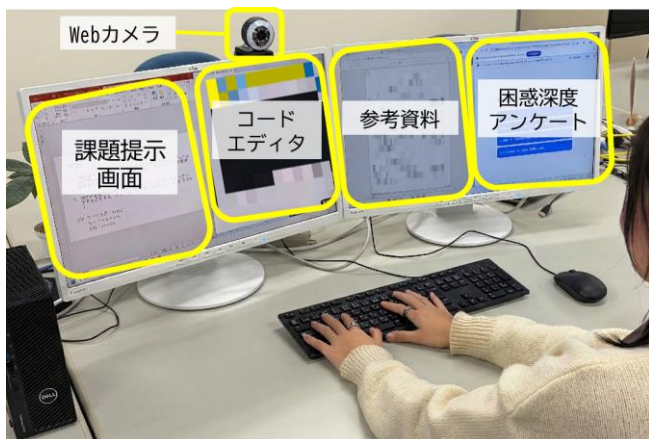


図3 実験環境

### 4.3. 実験手順

プログラミング課題は計2題実施し、設置したWebカメラを用い、表情画像を取得した。データセットの総数は126,125サンプルとなった。表情画像から深層学習であるTransformer[13]を用い、基本感情7分類スコアのデータセットを取得した。このデータセットに「3.2特徴量」にある特徴量生成を施し、説明変数とした。このとき式3ウィンドウサイズ $w$ は5とした。プログラミング課題と並行してWebアプリケーションによる4段階の困惑深度アンケートを実施した。本研究における「困惑」は教育心理学における「発達の最近接領域(ZPD)」の状態[14]と近いと考えた。ZPDでは学習支援の最適化において、学習者が

独力では解決に至らないものの、教育的介入によって達成可能となる心理領域を指す。ZPDを通過する学習過程は「自動化」、「自己による支援」、「有能な他者による支援」の段階があるとされている。そこで、本研究ではこの理論に基づき困惑深度の評価基準を設定した(表2)。アンケートは被験者が困惑を生じた瞬間にディスプレイ上のボタンを速やかに押下するよう指示した。ボタンが押下された時刻は、0.1秒の精度で自動的に記録する。押下され、次に押下されるまでの困惑状態は最初に押下された状態と定義した。以上の手順で取得した困惑深度アンケート結果を機械学習の目的変数とした。取得した説明変数と目的変数をタイムスタンプで結合し、困惑推定モデルを構築した。

表2 困惑深度評価基準

困惑深度	評価基準
0	非困惑(手が動いている状態)
1	少し困惑(手が止まり、想起する状態)
2	困惑(配布資料や教材で調べる状態)
3	とても困惑(他者に質問する状態)

### 4.4. 最適な分類器の調査

最適なモデルを選定するために全被験者モデルのデータセットを結合し、各機械学習モデル及び回帰分析における性能評価を実施した。本研究の考察に用いるモデルとして、最もF値が高い機械学習モデルを採用することとした。比較した学習モデルはRandom Forest(RF)[15]、Light Gradient Boosting Decision Tree(LightGBM)[16]、Gradient Boosting[17]、Support Vector Machine(SVM)[18]、AdaBoost[19]及び回帰分析(Logistic Regression)である。なお、各指標の算出にはクラス間のサンプル数の偏りを考慮し、マクロ平均を採用した結果を表3に示す。

表3 各モデルの評価結果

モデル	再現率	適合率	F値
RF	0.8897	0.8917	0.8882
LightGBM	0.7771	0.7787	0.7727
Gradient Boosting	0.6935	0.6929	0.6852
SVM	0.5948	0.5850	0.5726
AdaBoost	0.5689	0.5705	0.5632
Logistic Regression	0.6119	0.6032	0.6022

性能評価の結果、本研究ではF値が最も高いRFを最適な困惑推定モデルとして選定した。

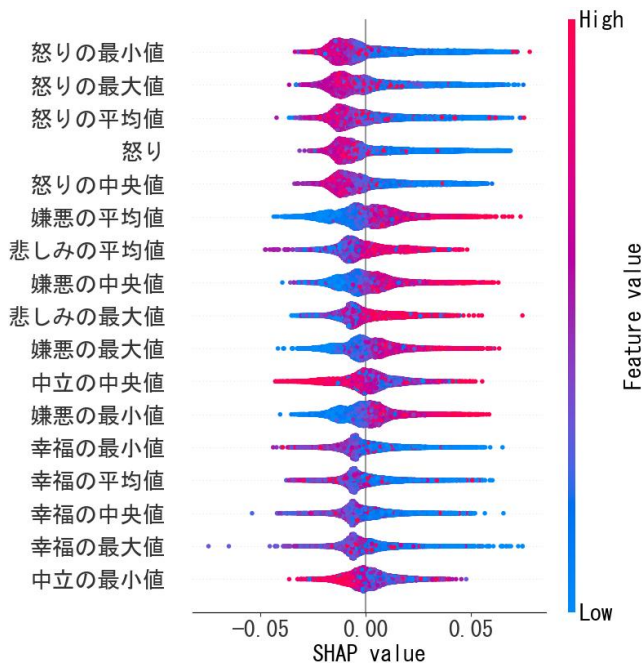
#### 4.5. 結果

教師付き機械学習 RF を用い困惑推定モデルを構築した。モデルの評価は F 値を使用し，基本感情複合分析は SHAP を使用した。

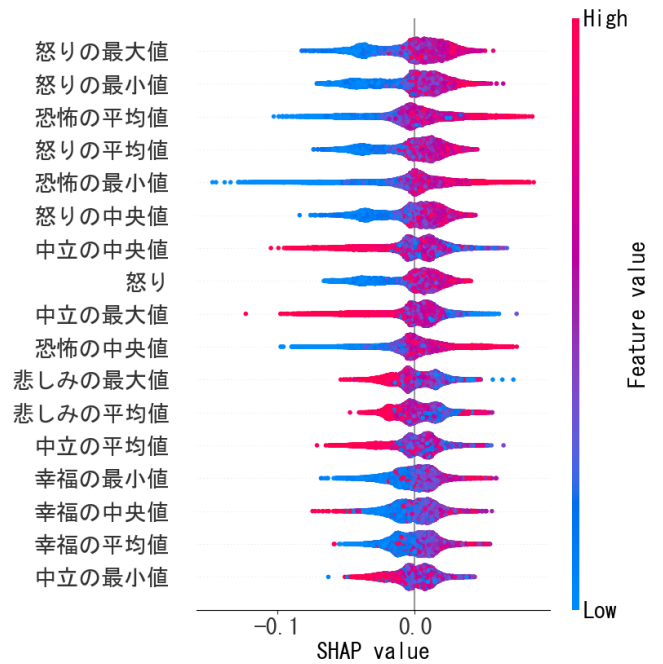
全被験者のデータセットを統合したときの混同行列を図 4 に各困惑深度における Violin Summary Plot を図 5 に示す。

True label \ Predicted label	0	1	2	3
0	7907	224	766	0
1	383	3858	590	0
2	358	125	10868	0
3	42	1	15	87

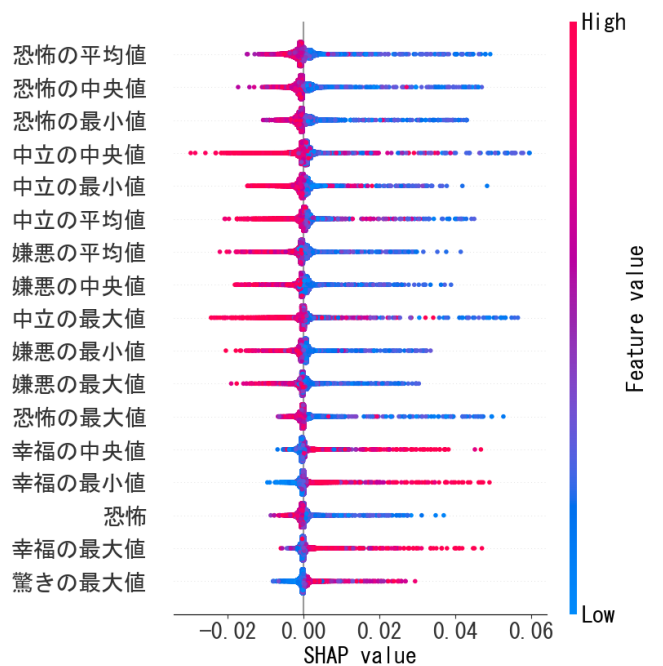
図 4 全被験者統合の混同行列



(a) 困惑深度 1



(b) 困惑深度 2



(c) 困惑深度 3

図 5 全被験者統合の Violin Summary Plot

図 5(a)から困惑深度 1 では「嫌悪」,「悲しみ」に関係する特徴量が正の方向に影響を与えることがわかった。このため,「嫌悪」,「悲しみ」の複合感情が少し困惑している状態を示唆すると考えられる。図 5(b)から困惑深度 2 では「怒り」,「恐怖」に関係する特徴量が正の方向に影響を与えることがわかった。このため,「怒り」,「恐怖」の複合感情が困惑状態を示唆すると

考えられる。図 5(c)から困惑深度 3 では「幸福」、「驚き」が正の方向に影響を与えることがわかった。このため、「幸福」、「驚き」の複合感情がとても困惑状態を示唆すると考えられる。

また、全被験者統合の SHAP 値と顕著に違いが見られた 2 名の被験者について示す。被験者 A の混同行列を図 6 に SHAP を図 7 に示し、被験者 B の混同行列を図 8 に SHAP を図 9 に示す。

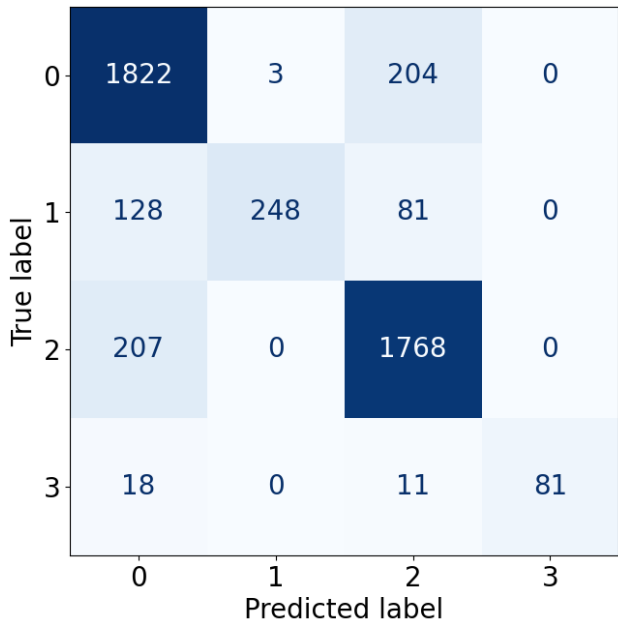
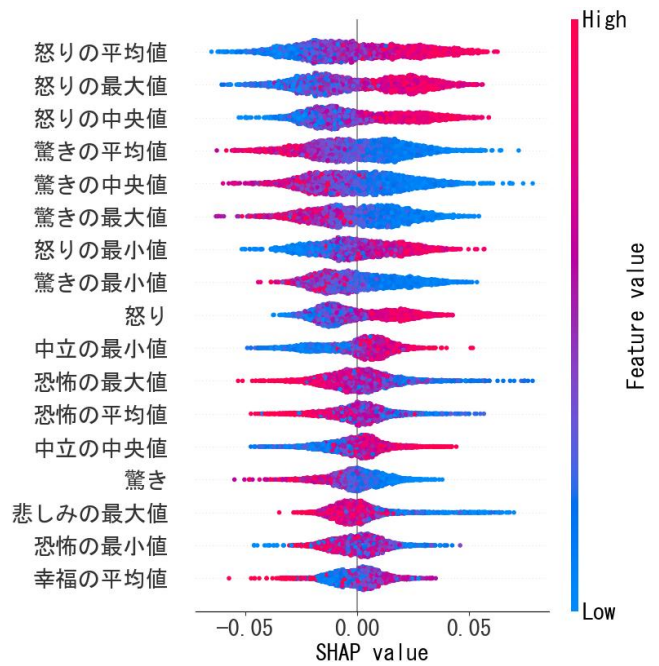
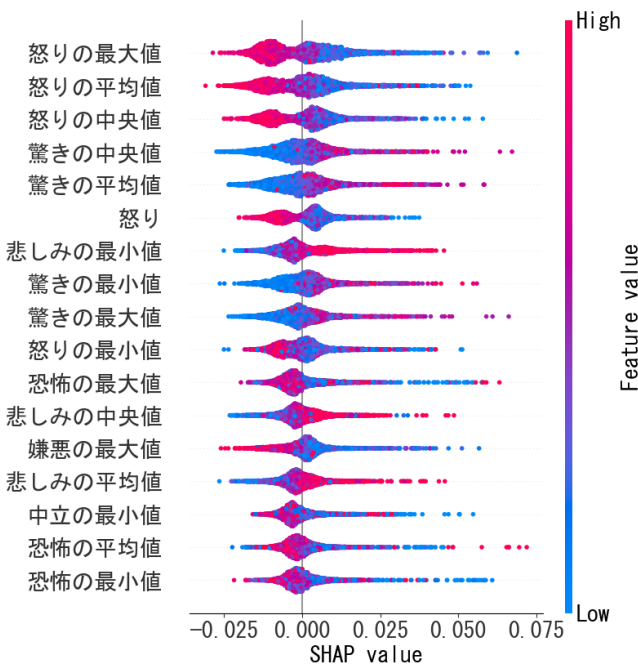


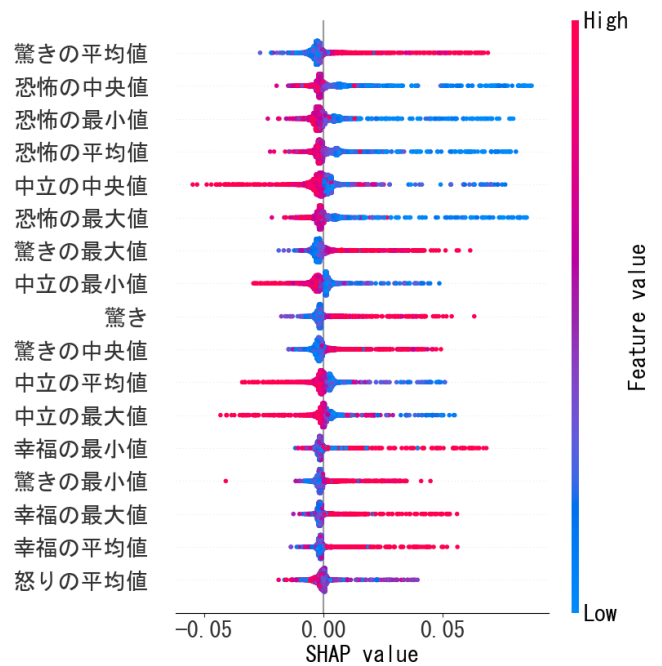
図 6 被験者 A の混同行列



(b) 困惑深度 2



(a) 困惑深度 1



(c) 困惑深度 3

図 7 被験者 A の Violin Summary Plot

図 7(a)から困惑深度 1 では「驚き」、「悲しみ」に関係する特徴量が正の方向に影響を与えることがわかった。このため、「驚き」、「悲しみ」の複合感情が少し困惑している状態を示唆すると考えられる。図 7(b)から困惑深度 2 では「怒り」、「中立」に関係する特徴量が正の方向に影響を与えることがわかった。このため、「怒り」、「中立」の複合感情が困惑状態を示唆すると

考えられる。図 7(c)から困惑深度 3 では「驚き」、「幸福」が正の方向に影響を与えることがわかった。このため、「驚き」、「幸福」の複合感情がとても困惑状態を示唆すると考えられる。

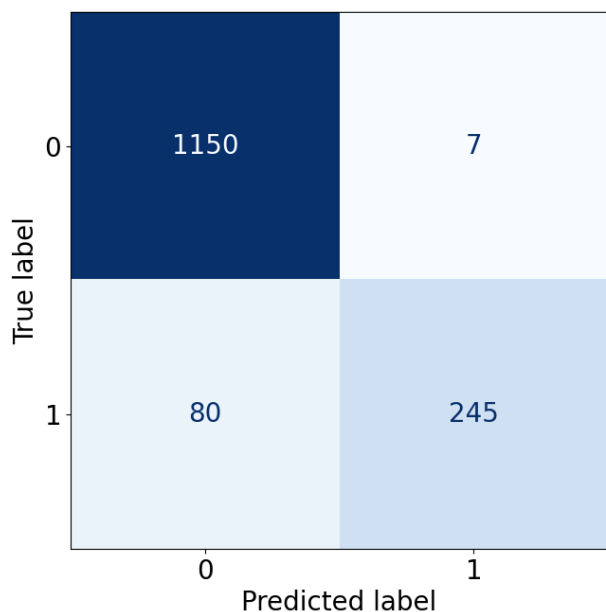


図 8 被験者 B の混同行列

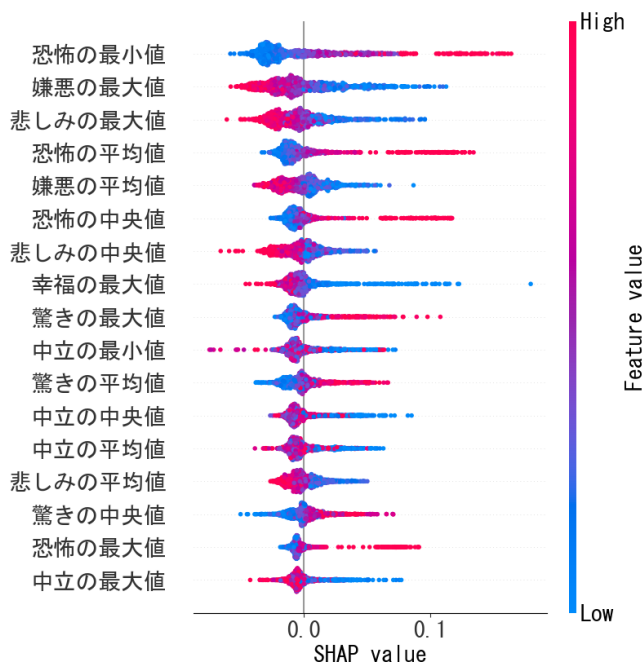


図 9 被験者 B の困惑深度 1 の Violin Summary Plot

図 9 から困惑深度 1 では「恐怖」、「驚き」に関する特徴量が正の方向に影響を与えることがわかった。このため、「恐怖」、「驚き」の複合感情が少し困惑している状態を示唆すると考えられる。また、他の困惑深度は実験中に検出できなかった。

混同行列から求められる各困惑深度における被験

者ごとの F 値を表 4 示す。

表 4 各困惑深度における被験者ごとの F 値

困惑深度	全被験者	被験者 A	被験者 B
0	0.899	0.867	0.964
1	0.854	0.701	0.849
2	0.921	0.876	
3	0.750	0.848	

## 5. 考察

### 5.1. 全体モデルと個人モデルの精度比較

表 4 より、全体モデルでは平均 0.856 の F 値となった。これは、困惑推定モデルにおいて高い精度となった。各困惑深度において、困惑深度 3 データ数が他の困惑深度と比べて少ないため学習が十分ではなく、F 値の低下がみられたと考えられる。

### 5.2. 自己による支援

図 5, 7, 9 から困惑深度 1 に着目すると、困惑に影響を与える特徴量に差異があることがわかった。全体モデルでは「嫌悪」と「悲しみ」の寄与度が高い。一方、個人モデルにおいて被験者 A では「驚き」と「悲しみ」、被験者 B では「恐怖」と「驚き」の寄与度が高い。被験者ごとに感情が異なるため、個人モデルの SHAP 値を確認することで困惑の把握をすることが可能である。

### 5.3. 有能な他者による支援

図 5 において、困惑深度 3 における「幸福」が正の方向に影響を与えていることがわかった。Russell[20]の感情の円環モデルにおいて「幸福」と「困惑」は、感情価の軸において対照的な位置に存在するとされている。これは、他者とのコミュニケーションによる安堵や笑顔から「幸福」が多く見られた。一方で、表 4 および図 7(c)より、被験者 A の困惑深度 3 が全体より 0.1 スコアが上昇した理由は困惑深度 3 において「驚き」という固有の感情が現れたためである。

### 5.4. 時間的変化の有効性

SHAP の図において重要な特徴量を見るとスライドウィンドウで統計的特徴量増加した特徴量が上位を占めていることがわかった。これは関連研究[10]が示す通り、困惑検出する際、一瞬の画像から得られる単なる分類モデルでは分類できず、機械学習を適用する際は時間的変化を特徴量の中に持たせ、分類することが必要であることが明らかになった。

### 5.5. 非対面学習における有効性

図 5, 7, 9 において、多くの特徴量において正の方

向または負の方向に赤い点が集中していることがわかった。ここから基本感情は表情表出が明確であると示唆している。顔の表情という日常的な情報から内面を可視化することで教育現場に心理的な安全性を保ちながら、データに基づいた教員が考案した個別指導を導入できる道筋を立てることができる。教育者に対して学習者の内面世界への深い洞察を与え、非対面学習における作業停滞の早期発見することが可能となる。

## 6. おわりに

本研究では、困惑は複数の基本感情の要素を含んでいることに着目し、基本感情の複合感情からプログラミング作業時の困惑を推定する手法を提案した。結果から、構築した困惑推定モデルは 0.8882 の F 値となった。また、SHAP を用いた分析から被験者毎に困惑推定が可能であること、時間的変化が重要であること、基本感情は表情から視認可能であることが明らかになった。しかし、個人モデルにおいて、データ数が極めて少ない困惑深度が存在し、不均衡データによる学習の偏りが生じていることが明らかになった。

今後は、学習データの質を担保するための実施課題の再選定および、被験者数の拡充によるデータセット全体の拡充を図る。また、スマートリング等の生体センシング機器を導入し、継続的かつ詳細なデータ収集を実現することも検討していく。

## 参 考 文 献

- [1] 総務省, “令和 6 年 通信利用動向調査書(企業編)”, [https://www.soumu.go.jp/johotsusintokei/statistics/pdf/HR202400\\_002.pdf](https://www.soumu.go.jp/johotsusintokei/statistics/pdf/HR202400_002.pdf), 2024 (参照 2026-01-05) .
- [2] 日本私立大学連盟, “私立大学における遠隔授業等の実施状況等に関する調査結果について”, [https://www.shidai.or.jp/topics\\_details/id=4151.pdf](https://www.shidai.or.jp/topics_details/id=4151.pdf), 2021 (参照 2026-01-05) .
- [3] ソニー損保, “Web 会議と安心に関する調査”, <https://prtimes.jp/a/?c=53388&r=8&f=d53388-8-1b2973afd75fa874698521938823dc45.pdf>, 2021 (参照 2026-01-05) .
- [4] 国立情報学研究所, “遠隔授業に関するアンケート調査の概要”, [https://www.nii.ac.jp/event/upload/20200914\\_Report.pdf](https://www.nii.ac.jp/event/upload/20200914_Report.pdf), 2020 (参照 2026-01-05) .
- [5] 布施泉, “初学者を主対象とする大学の一般プログラミング教育のオンライン授業による実施”, 高等教育ジャーナル 高等教育と生涯学習, Vol. 28, pp. 65-72, 2021.
- [6] P. Rozin and A. B. Cohen, “High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans”, *Emotion*, Vol. 3, pp. 68-75, 2003.
- [7] S. D'Mello and A. Graesser, “Dynamics of affective states during complex learning”, *Learning and Instruction*, Vol. 22, No. 2, pp. 145-157, 2012.
- [8] R. S. Baker, S. K. D'Mello, M. M. T. Rodrigo and A. C. Graesser, “Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments”, *International Journal of Human-Computer Studies*, Vol. 68, No. 4, pp. 223-241, 2010.
- [9] F. I. Yasser, B. H. Abd and S. M. Abbas, “Detection of confusion behavior using a facial expression based on different classification algorithms”, *Engineering and Technology Journal*, 2021.
- [10] M. Pachman, et al., “Eye tracking and early detection of confusion in digital learning environments: Proof of concept”, *Australasian Journal of Educational Technology*, Vol. 32, No. 6, 2016.
- [11] M. S. Hussain, R. A. Calvo and F. Chen, “Automatic Cognitive Load Detection from Face, Physiology, Task Performance and Fusion during Affective Interference”, *Interacting with Computers*, Vol. 26, pp. 256-268, 2014.
- [12] S. M. Lundberg and S. I. Lee, “A unified approach to interpreting model predictions”, *Proc. of NIPS 2017*, pp. 4765-4774, 2017.
- [13] A. Vaswani, et al., “Attention is all you need”, *Proc. of NIPS 2017*, pp. 5998-6008, 2017.
- [14] L. Cai, M. M. Msafiri and D. Kangwa, “Exploring the impact of integrating AI tools in higher education using the Zone of Proximal Development”, *Education and Information Technologies*, Vol. 30, No. 6, pp. 7191-7264, 2025.
- [15] L. Breiman, “Random forests”, *Machine Learning*, Vol. 45, No. 1, pp. 5-32, 2001.
- [16] G. Ke, et al., “Lightgbm: A highly efficient gradient boosting decision tree”, *Proc. of NIPS 2017*, pp. 3147-3157, 2017.
- [17] A. Natekin and A. Knoll, “Gradient boosting machines, a tutorial”, *Frontiers in Neurorobotics*, Vol. 7, p. 21, 2013.
- [18] S. Suthaharan, “Support Vector Machine”, *Machine Learning Models and Algorithms for Big Data Classification*, Springer, pp. 207-235, 2016.
- [19] R. E. Schapire, “Explaining Adaboost”, *Empirical Inference: Festschrift in Honor of Vladimir N. Vapnik*, Springer, pp. 37-52, 2013.
- [20] Russell, J. A.: A circumplex model of affect, *Journal of Personality and Social Psychology*, 39(6), pp.1161-1178 (1980).