

指パッチンのエンターテインメント性とリアルタイムジェスチャ識別の評価

永幡 峻太^{1,a)} 杉浦 裕太^{1,b)}

概要：指パッチンは、動作に伴って音および触覚フィードバックが自然に生じる身体ジェスチャであり、エンタテインメント入力としての活用が期待される。しかし、指パッチン入力のみならず体験的価値を、他の入力手法と比較して定量的に評価した研究は十分でない。また、リアルタイムアプリケーションに導入するには、入力の応答性を損なわずにジェスチャを安定して識別できる実装手法が求められる。そこで本研究では、1チャンネルEMGセンサを用いた指パッチン入力に着目し、リズムゲームタスクによる主観評価と、短信号長条件(8-256ms)における識別精度と遅延時間のトレードオフ評価を行った。10名の被験者による比較実験の結果、指パッチン操作はボタン操作よりも楽しさが有意に高かった。さらに、信号長を短縮した場合でも、例えば64ms条件において2クラス分類で平均70%以上の精度を維持することができた。

1. はじめに

近年、VR/ARをはじめとする拡張現実技術の普及に伴い、ユーザの身体動作を入力として用いるインタフェースが広く利用されている。中でも手指ジェスチャは、専用デバイスを必要とせず直感的に操作できることから、多様なアプリケーションで採用されている。

本研究では、日常のかつ瞬間的な身体動作である指パッチン(finger snapping)に着目する。指パッチンは、動作に伴って音や触覚フィードバックが自然に生じるという特徴を持ち、自己生成的なインタラクションとしてエンタテインメント用途に適した入力手段になり得る。

これまでも、指パッチンを入力として扱う研究は複数報告されており^{[1][2][3][4]}、検知・入力トリガとしての実用性は示されてきた。しかし、指パッチンという動作そのものがもたらす体験的価値、特にエンタテインメント文脈における楽しさを、他の入力手法と比較して定量的に評価した研究は十分ではない。

また、エンタテインメント入力として実際のアプリケーションに導入することを考えると、体験的価値に加えて、入力の応答性や認識の安定性といった実装上の要件も重要となる。入力動作が楽しいものであっても、入力が遅延したり、意図した動作が正しく認識されなかったりすると、体験全体の満足度は大きく損なわれる可能性がある。特に



図1 概要図

指パッチンのような瞬間的なジェスチャでは、短時間の信号から安定して入力を識別できるかどうか実用性を左右する。

そこで本研究の目的は、大きく二つある。一つ目は、指パッチン入力以外の入力手法と比較してどのような体験的価値を持つのかを明らかにすることである。特に、エンタテインメント文脈における「楽しさ」に着目し、リズムゲームタスクを用いた主観評価によって定量的に評価する。二つ目は、指パッチン入力を実際のリアルタイムアプリケーションに導入することを想定し、短信号長条件においても安定してジェスチャを識別可能であるか、またその際の遅延時間がどの程度となるかを評価することである。

¹ 慶應義塾大学

^{a)} r_nagahata@keio.jp

^{b)} sugiura@keio.jp

2. 関連研究

2.1 エンタテインメント体験の評価に関する研究

楽しさを含むユーザ体験の定量評価は HCI で広く議論されてきた [5][6][7]. 中でも内発的動機付け尺度 IMI[8] の興味・楽しさ下位尺度は, タスクやインタラクションの違いによる「楽しさ」を比較する目的で用いられることが多く, その信頼性が報告されている. 本研究では, 入力手法の違いに起因する体験差を比較するために IMI を参考にした質問紙を用い, 指パッチン入力の体験的特性を明らかにする.

2.2 ハンドジェスチャ識別研究

2.2.1 信号長短縮に着目した研究

信号長がジェスチャ識別性能に与える影響については, これまでも検討されている. Montazerin ら [9] や Ceolini ら [10] は, 多チャンネル EMG を用い, 窓長の増加に伴って識別精度が向上することを示した. 一方, Montecinos ら [11] は 1 チャンネル構成における結果も報告しているが, 最短の信号長は 500 ms に留まっている.

このように, 既存研究の多くは多チャンネル構成を前提とするか, 1 チャンネルであっても比較的長い信号長を対象としており, 300 ms 以下の短い信号長条件における少チャンネルでのリアルタイム性能 (識別精度と遅延時間) を体系的に評価した研究は十分ではない. 本研究は, 8–256 ms の短い信号長と 1 チャンネル EMG に焦点を当て, このトレードオフを明らかにする点に特徴がある.

2.2.2 1 チャンネル EMG デバイスを用いた研究

Arozi ら [12] は, BITalino を用いた 1 チャンネル EMG より 9 種類の手の動作を PCA+ANN でオフライン分類し, 単一チャンネルでも複数動作を識別可能であることを示した. 一方, 同研究では信号長短縮時の性能変化やリアルタイム処理の応答遅延は検討されていない. 本研究はこの構成を踏まえ, 短信号長条件でのリアルタイム識別性能と遅延時間評価へと拡張する.

3. エンタテインメント体験の評価

本章では, 指パッチン入力のエンタテインメント体験を, 他入力手法と比較して評価するために作成したリズムゲームタスクの設計と実装, および主観評価結果と考察を述べる.

3.1 リズムゲームタスク

3.1.1 タスク設計と入力条件

筋活動に基づく入力, またはキーボード入力によってノーツを消去できるリズムゲームタスクを作成した. 入力条件は, ボタン操作, 指パッチン操作, ピンチ操作, グリップ操作, 空中タップ操作の 5 条件である. 従来のリズム

表 1 アンケート項目 (楽しさの評価)

No.	質問項目
Q1	ゲーム体験はとても楽しかった.
Q2	もう一度ゲームをプレイしたいと感じた.
Q3	ゲーム体験は退屈だと感じた.
Q4	ゲーム体験をしている間「楽しいな」とずっと考えていた.
Q5	ゲーム体験に没頭できなかった.

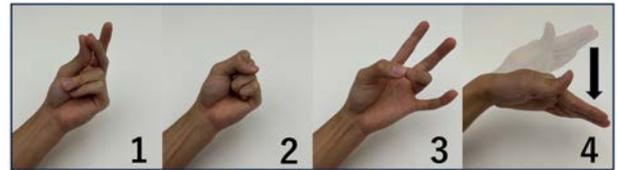


図 2 各ジェスチャの様子

ムゲームで一般的なボタン操作に加え, 比較対象として複数ジェスチャを含めることで, 指パッチン入力の体験的特性を相対的に評価する.

3.1.2 キャリブレーションとタイミング補正

タスク実施前に, 各ジェスチャ条件ごとに閾値キャリブレーションを行った. 誤検知を防ぎつつ筋活動を確実に検出することを目的として, 筋電位信号の振幅を基に被験者ごとに閾値を手動設定した. さらに, 入力手法間でノーツ消去の提示タイミングを揃えるため, 閾値設定とは独立にタイミング調整を行った. 具体的には, BPM60 で等間隔に流れる 10 個のノーツを各ジェスチャで消去してもらい, 実際のノーツ消去時刻と理想的なビートポイントとの差から遅延時間を算出し, その値に基づいて補正を適用した. これにより, 入力手法間で公平に比較できる実験環境を構築した.

3.1.3 実験手順と主観評価

実験は 2 日間に分けて実施した. 1 日目はボタン操作と指パッチン操作, 2 日目は指パッチン操作とその他ジェスチャ (ピンチ, グリップ, 空中タップ) を実施した. なお, 指パッチンは比較の基準条件として両日で共通に実施した. 各ジェスチャの様子を図 2 に示す.

各条件終了後に, IMI の興味・楽しさ下位尺度を参考に作成した質問紙 (表 1) に 7 段階 Likert 尺度で回答を求めた. 否定的項目は反転し, 合計スコアを楽しさ指標として算出した. また, 楽しさ以外の体験的側面 (疲労感, 難しさ, 直感性, ストレス等) も同様に 7 段階で評価し, 入力手法間の差を多角的に分析した.

3.2 リズムゲームタスクの実装

3.2.1 ゲームフロー

リズムゲームタスクは Unity 上で構築した. 画面左右いずれかからビートポイントに向かって一定速度 (BPM60) でノーツが移動し, 被験者は指定入力でノーツを消去する. 筋活動入力を用いる条件では, 電極を装着した手と反対側

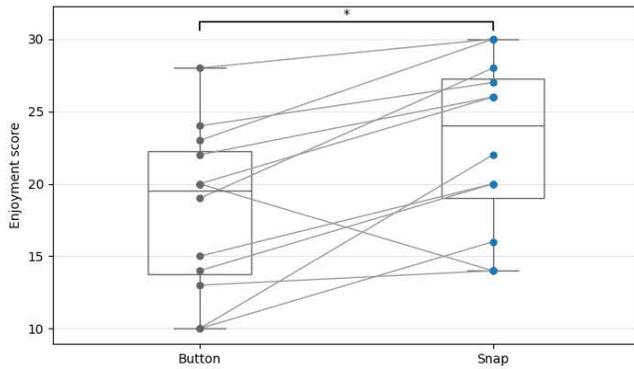


図 3 楽しさのスコア

表 2 リズムゲームタスクの性能評価結果

評価項目	値
閾値検知からノーツ消去までのタイムラグ [ms]	13.55
タイミングキャリブレーション時の時間差 [ms]	89.56
誤検知率 (False Positive)	0.024
未検知率 (False Negative)	0.021

からノーツが出現するよう設定し、視覚的バイアスを抑制した。

3.2.2 閾値検知と UDP 通信

筋活動入力では、Python 側で閾値超過を検知し、そのイベントを UDP パケットとして Unity に送信した。Unity 側は受信イベントに基づきノーツ判定を行う。この分離により、筋活動検知処理と描画処理を独立させ、安定したリアルタイム動作を実現した。

3.3 主観評価結果

3.3.1 リズムゲームタスクの参加者

10 名の被験者 (20.3 ± 0.7 歳 (SD)) を対象にリズムゲームタスクを実施した。

3.3.2 タスクの評価

本研究で用いるリズムゲームタスクが入力手法の比較に適した実験タスクであるかを確認するため、タスク自体の性能評価を行った。評価指標として、EMG 閾値超過の検知からノーツ消去までのタイムラグ、タイミングキャリブレーション時における理想ビートポイントと実際消去時刻の差、誤検知率 (False Positive)、未検知率 (False Negative) を用いた。誤検知率はノーツ消去に関係しない入力回数/総入力回数、未検知率は消去されなかったビート数/総ビート数として算出した。結果を表 2 に示す。

3.3.3 楽しさの評価

入力手段ごとの楽しさスコアを図 3 に示す。ボタン操作と指パッチン操作の比較について Wilcoxon 符号付順位検定を行った結果、指パッチン操作はボタン操作よりも有意に高い楽しさを示した ($p = 0.012$)。また、指パッチン・ピンチ・グリップ・空中タップの 4 条件比較では、Friedman 検定の結果、有意差は認められなかった ($p = 0.95$)。ただ

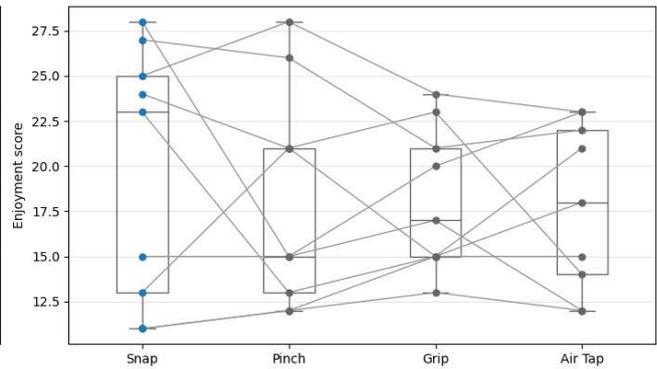


表 3 楽しさ以外の主観評価結果 (ボタン操作と指パッチン)

項目	ボタン	指パッチン	Wilcoxon p
気持ちがよかった	4.00	4.33	0.570
気に入った	3.25	4.25	0.151
タイミングが難しかった	2.67	3.83	0.0269
リズムが取りやすかった	5.75	4.33	0.0242
自然に使えた	6.33	5.00	0.0269
ストレスに感じた	1.33	3.67	0.0101
直感的にわかりやすかった	6.25	5.33	0.196
疲れた	1.50	4.75	0.0044
操作の実感が湧いた	4.17	5.25	0.123
楽しめた	3.08	4.67	0.0371

し平均値は指パッチン操作が最も高かった。

3.3.4 楽しさ以外の指標評価

ボタン操作と指パッチン操作における、楽しさ以外の主観評価結果を表 3 に示す。有意差が認められた項目については、尺度の方向に基づき、望ましい方向のスコアを太字で示した。

次に、4 種類のジェスチャ条件で楽しさ以外の主観評価指標を比較した。Friedman 検定の結果、「疲れた」という項目に対して条件間で有意差が認められた ($p = 0.0288$)。指パッチン操作は他条件と比較して疲労感が高い傾向を示した。

3.4 考察

主観評価の結果、指パッチン操作はボタン操作よりも楽しさのスコアが高かった。これは、指パッチンが音と触覚フィードバックを伴う自己生成的なインタラクションであり、操作実感や達成感を強めた可能性がある。一方で、指パッチン操作は「疲れた」「ストレスに感じた」「タイミングが難しかった」がボタンより高く、操作コストが増大していることが示された。すなわち、指パッチン入力は楽しさの向上と引き換えに、疲労や難しさを増やすトレードオフを持つ入力手段である。

この結果から、エンタテインメント用途で指パッチン入力を設計する際の含意として、以下が考えられる。

- **短時間・高揚感のある場面に適する:** 楽しさの利点を活かすには、長時間連続入力を要求しない設計 (短い

セッション、入力頻度の制御)が有効である。

- **誤差許容とフィードバック設計:** タイミングの難しさが増えるため、判定の許容幅や補助フィードバック(視覚・聴覚)を調整することで負荷を下げられる可能性がある。
- **適用範囲の限定:** 4条件比較で楽しさに有意差が出なかった点を踏まえると、常に指パッチンが最良とは限らない。タスクの目的(爽快感・没入・達成感)に応じてジェスチャを選択すべきである。

4. 識別精度と遅延時間

本章では、指パッチンを含む複数ジェスチャを実時間で識別するための手法と、短信号長条件(8-256ms)における識別精度と遅延時間の関係を示す。

4.1 識別用データ収集

ジェスチャ識別に用いるデータとして、指パッチン(Snap)、ピンチ(Pinch)、グリップ(Grip)、空中タップ(Air Tap)の4種類を対象にEMGデータを収集した。データ収集は被験者ごとに1分間のタスクとして実施し、タスク中では3秒ごとに指定ジェスチャを繰り返し実行する形式とした。この1分間タスクを、指パッチンは2回、その他は各1回ずつ実施した。実験条件の統一のため、電極は被験者ごとに新しいものを使用した。

4.2 計測環境および識別処理

4.2.1 ハードウェアおよび計測環境

筋電位計測デバイスとしてBITalino(Creact社)を用いた。サンプリングレートは1000Hzに設定し、取得したEMGデータはBluetooth通信を介してノートパソコンへ送信した。実験にはASUS Zenbook 14 OLED(UX3405MA-U9321W)を使用し、データ取得・処理はPythonで実装した。

4.2.2 前処理および信号区間の抽出

リアルタイム処理では、EMG信号が閾値を初めて超過した時刻を起点として、そこから一定長の信号区間を抽出した。抽出信号長は8ms, 16ms, 32ms, 64ms, 128ms, 256msの6条件とした。これらはFFTを効率的に計算可能なよう、サンプル数が2の累乗となるように設定している。

閾値は、各被験者の事前データ収集区間に対してピーク振幅を算出し、誤検知が生じない範囲で筋活動を確実に検出できる値として手動で決定した。オンライン処理では、この閾値を固定して用いた。

4.2.3 特徴量抽出および識別モデル

抽出区間から時間領域および周波数領域の特徴量を算出し、合計21次元の特徴ベクトルを構成した(表4)。時間領域特徴量としてIEMG, MAV, RMS, 波形長等を、周

表4 抽出した特徴量一覧

種類	特徴量名	定義
時間領域	IEMG	絶対値の総和
	MAV	絶対値の平均
	MAV1	信号前半における絶対値の平均
	MAV2	信号後半における絶対値の平均
	SSI	二乗値の総和
	VAR	分散
	RMS	二乗平均平方根
	WL	波形長(隣接するサンプル間の差分絶対値の総和)
	DASDV	差分の二乗平均平方根
	AR1	自己回帰係数(次数1)
	AR2	自己回帰係数(次数2)
	AR3	自己回帰係数(次数3)
	AR4	自己回帰係数(次数4)
	Hjorth1	活動性(信号の分散)
Hjorth2	移動性(一次差分と元信号の分散比の平方根)	
Hjorth3	複雑度(一次差分と二次差分の分散比)	
周波数領域	FFT_max_amp	振幅スペクトルの最大値
	FFT_mean_amp	振幅スペクトルの平均値
	FFT_median_amp	振幅スペクトルの中央値
	FFT_centroid	スペクトル重心
	FFT_energy	振幅の二乗和

波数領域特徴量としてFFTに基づく指標を用いた。

識別モデルには全結合型の人工ニューラルネットワーク(ANN)を用いた。特徴量に対しPCAを適用し、第3主成分までをANN入力とした。PCAは学習データのみに対して算出し、テストデータには学習データから得た変換を適用することで情報漏洩を防いだ。ANNの基本方針はAroziら^[12]に基づく。

4.3 遅延時間計測

短信号長条件における実時間性能を評価するため、オンラインで識別処理を実行した。本研究では、閾値超過の検知から、所定長の信号区間取得および特徴量抽出・PCA変換・ANN推論を経て識別結果を出力するまでの時間を遅延時間として計測した。遅延時間結果を表5に示す。ここで「識別までにかかった時間」は、閾値超過検知時刻から識別結果の出力までの総時間であり、信号区間の取得時間(窓長)と取得後処理時間の和として解釈できる。全条件で取得後処理時間は概ね10ms前後であり、総遅延時間の増加は主として窓長に起因していた。

4.4 識別性能評価結果

本節では、個人モデルの基本性能を確認した後、短信号長条件における識別精度と遅延時間の関係を示す。

表 5 信号長ごとの遅延時間および処理時間

取り出すデータ長 [ms]	256	128	64	32	16	8
識別までにかかった時間 [ms]	274.58	144.78	82.07	53.07	38.68	27.71
処理時間 [ms]	8.97	8.78	10.35	9.63	10.66	9.32

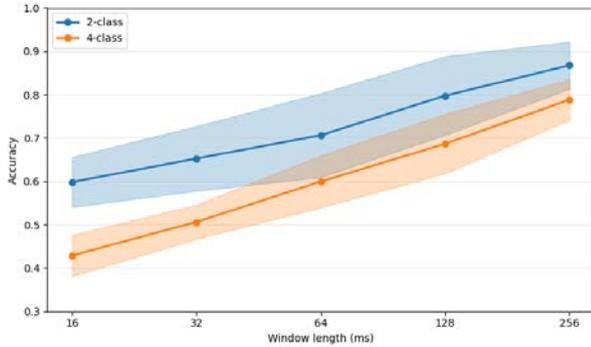


図 4 信号長ごとの識別精度

4.4.1 個人モデルにおける識別精度

本項では、後続の評価の前提として個人モデルの識別性能を確認する。識別タスクとして、指パッチン動作のみを検知する2クラス分類 (snap vs. other) と、4種類のジェスチャを区別する多クラス分類 (Snap/Pinch/Grip/Air Tap) を設定した。

2クラス分類では各クラス 40 サンプル (合計 80)、多クラス分類では各クラス 20 サンプル (合計 80) を用い、5分割交差検証により正解率 (accuracy) を算出した。本研究では、10人の被験者 (20.3 ± 0.7 歳 (SD)) から識別用のジェスチャデータを収集し、各被験者ごとに個人モデルを構築して識別精度を評価した。

4.4.2 信号長短縮に伴う識別精度と遅延時間のトレードオフ

信号長を 8 ms, 16 ms, 32 ms, 64 ms, 128 ms, 256 ms の6条件とし、各条件で識別精度および遅延時間を計測した。図4に信号長ごとの識別精度を示す。実線は被験者間平均の識別精度を示し、薄色の帯は被験者間の95%信頼区間を表す。信号長の短縮に伴い識別精度は低下する傾向が見られた。256 msにおいて最も高い識別精度が得られ、2クラス分類では最大93.8%、平均86.8%、4クラス分類では最大93.8%、平均78.8%であった。一方、64 ms条件においても2クラス分類では平均70%以上の精度が維持された。

5. 考察

5.1 識別精度と信号長の関係

識別精度は信号長が長いほど高くなる傾向を示した。これは、長い信号区間ほど筋活動の時間変化や周波数特性をより多く含み、ジェスチャ間の差異を捉えやすくなるためと考えられる。一方で、64 ms条件でも2クラス分類では平均70%以上を維持しており、指パッチン検知 (snap vs. other) のようにタスクを限定すれば、短い信号長でも実用

的な精度が得られる可能性を示唆する。

5.2 リアルタイムジェスチャ識別における遅延特性

遅延時間計測の結果、総遅延時間は主として信号区間の取得時間 (窓長) に依存し、取得後の処理時間 (特徴量抽出・PCA・ANN推論) は全条件で概ね10 ms前後に留まった。これは、本研究の処理パイプラインが窓長に対して計算量の増加が小さく、リアルタイム処理に適した構成であることを示している。従って、応答性を改善する最も直接的な手段は窓長短縮であるが、その際に精度低下が生じるため、アプリケーション要件に応じた窓長選択が必要となる。

本研究のリズムゲームタスクでは、タイミングキャリブレーションにより平均89.56 msの補正を適用しており、体験上は入力手法間の公平性を確保できた。この条件を踏まえると、64–128 msは「精度と応答性の両立」を狙う現実的な選択肢となり得る。

6. 今後の課題

本研究にはいくつかの限界がある。第一に、閾値は被験者ごとに手動で設定しており、運用容易性と再現性の観点で改善の余地がある。今後、筋活動の統計的特性に基づく自動閾値決定 (例: 安静時ノイズ分布に基づく閾値設定) を導入することで、被験者間のばらつきを抑制し、実運用に近い条件での評価が可能になると考えられる。

第二に、本研究の主観評価はリズムゲームタスクに限定されており、他のエンタテインメント文脈 (例: 音楽制作, VR体験, カジュアル操作) で同様の傾向が得られるかは未検証である。今後はタスク種類や入力頻度の異なる条件で検証し、指パッチン入力の適用範囲と設計要件を明らかにする必要がある。

第三に、被験者数が限られており ($n = 10$)、個人差 (指パッチンの得意・不得意, 筋活動特性) の影響を十分に分解できていない可能性がある。今後は被験者数を増やし、事前に得意度を測定した上で層別するなど、個人差を考慮した分析を行うことで、誰にとってどの入力が有効かをより精緻に議論したい。

参考文献

- [1] Hosseini, M., Mueller, H. and Boll, S.: Controlling the rooms: How people prefer using gestures to control their smart homes, *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp. 1–18 (2024).

- [2] Chen, M., Lin, J., Li, W., Zou, Y. and Wu, K.: SnapUnlock: A Contrastive Learning-Based Contactless Authentication via Heterogeneous Sensors, *Wireless Communications and Mobile Computing*, Vol. 2022, No. 1, p. 2134493 (online), DOI: <https://doi.org/10.1155/2022/2134493> (2022).
- [3] Yeh, Y.-T. and Vogel, D.: Snap! snap! SNAP! Acoustic Finger Snapping Ability for Interaction, *Proceedings of Graphics Interface (GI '25)*, Kelowna, BC, Canada, Canadian Human-Computer Communications Society (CHCCS), pp. ??-?? (2025).
- [4] Pomykalski, P., Woźniak, M. P., Woźniak, P. W., Grudzień, K., Zhao, S. and Romanowski, A.: Considering Wake Gestures for Smart Assistant Use, *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA '20, New York, NY, USA, Association for Computing Machinery, p. 1-8 (online), DOI: [10.1145/3334480.3383089](https://doi.org/10.1145/3334480.3383089) (2020).
- [5] van der Kooij, K., van Dijksseldonk, R., van Veen, M., Steenbrink, F., de Weerd, C. and Overvliet, K. E.: Gamification as a Sustainable Source of Enjoyment During Balance and Gait Exercises, *Frontiers in Psychology*, Vol. Volume 10 - 2019 (online), DOI: [10.3389/fpsyg.2019.00294](https://doi.org/10.3389/fpsyg.2019.00294) (2019).
- [6] Tate, Lyons, Kalyararaman, Ribisl, Ward and Bowling: Energy Expenditure and Enjoyment during Video Game Play: Differences by Game Type, (online), DOI: [10.17615/ereb-mp36](https://doi.org/10.17615/ereb-mp36) (2011).
- [7] Al-Khiami, M. I., Jaeger, M. and Soleimani, S. M.: The Integration of Digital Techniques in Engineering Education: A Case Study to Evaluate Student's Motivation and Performance, *International Journal of Information and Education Technology*, Vol. 13, pp. 844-848 (online), DOI: [10.18178/ijiet.2023.13.5.1877](https://doi.org/10.18178/ijiet.2023.13.5.1877) (2023).
- [8] Ryan, R.: Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory, *Journal of Personality and Social Psychology*, Vol. 43, pp. 450-461 (online), DOI: [10.1037/0022-3514.43.3.450](https://doi.org/10.1037/0022-3514.43.3.450) (1982).
- [9] Montazerin, M., Rahimian, E., Naderkhani, F., Atashzar, S. F., Yanushkevich, S. and Mohammadi, A.: Transformer-based Hand Gesture Recognition via High-Density EMG Signals: From Instantaneous Recognition to Fusion of Motor Unit Spike Trains (2022).
- [10] Ceolini, E., Frenkel, C., Shrestha, S. B., Taverni, G., Khacef, L., Payvand, M. and Donati, E.: Hand-Gesture Recognition Based on EMG and Event-Based Camera Sensor Fusion: A Benchmark in Neuromorphic Computing, *Frontiers in Neuroscience*, Vol. Volume 14 - 2020 (online), DOI: [10.3389/fnins.2020.00637](https://doi.org/10.3389/fnins.2020.00637) (2020).
- [11] Montecinos, C., Espinoza, J., Zamora Zapata, M., Meruane, V. and Fernandez, R.: Improving Fast EMG Classification for Hand Gesture Recognition: A Comprehensive Analysis of Temporal, Spatial, and Algorithm Configurations for Healthy and Post-Stroke Subjects, *Sensors*, Vol. 25, No. 22 (online), DOI: [10.3390/s25226980](https://doi.org/10.3390/s25226980) (2025).
- [12] Arozi, M., Caesarendra, W., Ariyanto, M., Munadi, M., Setiawan, J. D. and Glowacz, A.: Pattern Recognition of Single-Channel sEMG Signal Using PCA and ANN Method to Classify Nine Hand Movements, *Symmetry*, Vol. 12, No. 4 (online), DOI: [10.3390/sym12040541](https://doi.org/10.3390/sym12040541) (2020).