

畳み込みニューラルネットワークを利用した
Oculofeedback に関する研究

2020 年 3 月

博士 (工学)

惠本 序珠亜

中部大学

2019 年度

中部大学大学院工学研究科情報工学専攻

博士学位論文

畳み込みニューラルネットワークを利用した
Oculofeedback に関する研究

惠本 序珠亜

目次

第 1 章	序論	1
1.1	背景	1
1.2	目的	2
1.3	構成	3
第 2 章	Oculofeedback システム	4
2.1	Biofeedback	4
2.2	Neurofeedback	5
2.3	Decoded Neurofeedback	6
2.4	Oculofeedback の提案	6
	2.4.1 Oculofeedback システム	7
	2.4.2 対応機器	8
	2.4.3 プログラム	9
2.5	注意と眼球運動	12
	2.5.1 選択的注意 (Selective Attention)	12
	2.5.2 トップダウン注意 (Top-down Attention) と ボトムアップ注意 (Bottom-up Attention)	14
	2.5.3 顕在的注意 (Overt Attention) と 潜在的注意 (Covert Attention)	15
	2.5.4 注意の評価方法	15

2.5.5	注意に関わる眼球運動	16
2.5.6	マイクロサッカードに反映される注意	19
2.6	まとめ	19
第 3 章	Oculofeedback のための実時間高精度マイクロサッカード 検出	20
3.1	マイクロサッカードに基づく注意状態の推定	20
3.2	眼球運動の計測とマイクロサッカード検出	21
3.3	従来法	22
3.4	提案法	24
3.4.1	ConvNet の構造	24
3.4.2	学習方法	27
3.5	マイクロサッカード検出法の評価	29
3.5.1	マイクロサッカードデータセット	30
3.5.2	検出精度の定義	35
3.5.3	従来法の評価	36
3.5.4	提案法の評価	38
3.6	考察	41
第 4 章	Oculofeedback による注意状態操作実験	45
4.1	マイクロサッカードフィードバック実験	45
4.1.1	Oculofeedback による注意操作の仕組み	45
4.2	目的	46
4.3	方法	47
4.3.1	結果	49
4.3.2	考察	53
第 5 章	畳み込みニューラルネットワークを応用した新たな眼球運 動測定システム	55

5.1	Oculofeedback に適した眼球運動計測装置	56
5.2	提案法	57
5.2.1	概要	57
5.2.2	ConvNet の学習	59
5.2.3	モデルの構造	60
5.3	評価方法	63
5.3.1	眼球運動計測装置	63
5.3.2	手順	64
5.4	データセット	65
5.4.1	実験	65
5.4.2	データの抽出	66
5.4.3	アノテーション	68
5.5	評価結果	70
5.5.1	特徴点検出	70
5.5.2	イベント検出	73
5.6	考察	75
5.6.1	特徴点検出	75
5.6.2	イベント検出	76
5.6.3	精度と軽量化のトレードオフ	77
5.6.4	学習手順	77
5.7	まとめ	78
第 6 章	総括	80
6.1	本論文の成果	80
6.2	Oculofeedback の今後の展望	82
参考文献		85

第 1 章

序論

1.1 背景

眼から得られる情報は、様々な分野への応用が期待できる魅力的なデータである。

眼球運動には脳の広範な領域が関係しており [1, 2]，これらの領域の情報が直接・間接的に眼球運動に反映されると考えられている．そのため，眼球運動は脳の状態を客観的に評価する指標として，医療や教育，スポーツ，マーケティング，エンターテインメント，自動車安全技術など多様な応用研究の対象とされている [3, 4, 5, 6, 7, 8, 9]．既に，眼球運動から注意 [10] や認知 [11]，集中度 [12]，眠気 [13, 14] など，様々な脳の状態推定が試みられており，こうした応用研究では，脳の高次機能や生理的な状態を客観的に評価することで，これまで困難であった個人に合わせたサービスや人の心身に合わせた技術の開発を目指している．

近年こうした研究や実用場面で広く用いられている眼球運動計測装置は，カメラで撮影した眼の画像から，画像処理によって視線方向や瞳孔径を推定する．このような計測方式では，カメラと演算装置があれば非侵襲に眼球運動を計測できるため，例えば，監視カメラで撮影した人の顔画像から視線を

推定したり [15], 頭部に固定したカメラから高精度に眼球運動を計測したりするなど [16], 利用目的に合わせた実装が可能である. また, 脳波や心拍などのデータは身体にセンサーを取り付けて計測する必要があるが, 眼球運動データはセンサーを取り付けずに遠隔で計測可能である. 眼球運動データは生体情報の中でも計測対象の負担が少ない生体情報であると言える.

本研究では, 眼球運動データを脳の訓練に応用する手法として Oculofeedback を提案する. Oculofeedback は, 被験者から計測した眼球運動データに基づいて, 適切な感覚フィードバックを与えることにより, 被験者の脳の状態を操作する方法を提供する. 被験者が Oculofeedback による訓練に取り組むことにより, 眼球運動と相互に関係し合う注意や認知能力を向上させることが期待できる. こうした Oculofeedback には, 眼球運動から注意などの脳の状態をロバストに推定するための高度な眼球運動計測・解析アルゴリズムが必要である. 一般に, 生体情報に基づくヒトの状態推定は非定常なノイズや個人差の影響により難しい問題となる. これに対し, 近年, 生体情報の特徴抽出にディープラーニングを応用する研究が進んでいる [17]. 本研究では, Oculofeedback のような, 眼球運動からのロバストかつ実時間での特徴量推定が必要な手法のために開発した, ディープラーニングを応用した眼球画像解析法も提案する.

1.2 目的

本研究の目的は, Oculofeedback という新しい脳状態操作手法を提案し, その実現可能性を実証することである. そのために, Oculofeedback を実行するシステムを開発し, 例として, 注意の状態を操作する実験を実施する. また, 注意状態を操作する Oculofeedback で必要となる特定の眼球運動特徴量を抽出するために, 畳み込みニューラルネットワークを用いた解析法を開発し, その有効性を示す. さらに, 今後 Oculofeedback 技術の中核を担うものと予想される畳み込みニューラルネットワークで構成した新たな眼球画

像解析法について述べ、その性能の評価結果を示す。

1.3 構成

本論文は全 6 章で構成される。

第 1 章では本研究の背景と目的を述べた。

第 2 章では、本研究で提案する Oculofeedback について詳述し、Oculofeedback を具現化するために開発したシステムについて述べる。また、本研究で Oculofeedback の一例として実施した実験において扱う注意と眼球運動についてもここで述べる。

第 3 章では、注意状態を操作する Oculofeedback で用いるマイクロサッカードと呼ばれる微小眼球運動を、高精度に検出するために開発した手法について述べる。

第 4 章では、開発した Oculofeedback システムとマイクロサッカード検出法を用い、注意状態を操作する実験を実施した結果について述べる。

第 5 章では、眼球運動から注意状態などを推定するために必要な特徴量を抽出する、畳み込みニューラルネットワークを応用した眼球画像解析法について述べる。

最後に、第 6 章で本研究について総括する。

第 2 章

Oculofeedback システム

本章では，適切な眼の動きを促進することにより，脳の状態を操作することを可能にする Oculofeedback と名付けた新しいヒトの状態制御法を提案する．Oculofeedback は，眼球運動に反映される注意などの脳の状態を，被験者にフィードバックすることで，脳の状態を改善する訓練手法とも捉えられる．

本章では，Oculofeedback の基となる Biofeedback および Neurofeedback について説明し，Oculofeedback とその実施に必要なシステムとして開発した Oculofeedback システムについて述べる．

2.1 Biofeedback

Biofeedback は，随意的にコントロールすることが難しい生体状態を随意的にコントロールする方法を被験者に学習させる訓練手法である [18]．Biofeedback の歴史は長く，その概念は 100 年以上前には発案されており，北米の Biofeedback の協会は 1969 年に成立している [19]．Biofeedback では，被験者が通常自覚することが難しい生体状態を計測機器と状態推定アルゴリズムからなる何らかのシステムにより推定し，被験者へ知覚可能な視聴

覚等の感覚刺激でフィードバックする．これにより，被験者はどのような意識が生体状態の改善に結びつくのかを学習する．

たとえば，緊張性頭痛の患者の場合は，筋肉の緊張を和らげるために，筋電図 (EMG) から得られる筋緊張状態を反映する特徴量をフィードバックする．被験者はフィードバックにより筋緊張の状態を知ることができ，自身の筋緊張が和らいでいる状態をどのようにすれば維持できるかを学習することで頭痛を軽減する [20]．

リハビリテーション医療では，このような EMG を使った Biofeedback が幅広い疾患に対して応用されており，標準的な治療方法となりつつある [21]．また，脳卒中の治療ガイドラインでは，歩行障害のリハビリテーションに，EMG や関節角度を用いた Biofeedback を推奨している [22]．従って Biofeedback は，その有効性が広く認められた手法であると言える．

2.2 Neurofeedback

Biofeedback の中でも脳波 (EEG) などの脳からの計測データを利用したものは，Neurofeedback と呼ばれる．Neurofeedback の臨床応用の中でも，注意欠陥・多動性障害 (ADHD) に対する効果が注目されている [23]．EEG を利用した Neurofeedback では，例えば，ADHD を改善するために，被験者にディスプレイやスピーカーで視聴覚刺激を与えることにより，EEG で計測される脳の活動状態を，ADHD の発生していないより望ましい状態に近づける．このとき，被験者に，現在どれだけ望ましい状態に近づいているかを知らせるフィードバックを与え，被験者に ADHD を抑える感覚を学習させる．このような訓練を繰り返すことにより，被験者が自身の意識により注意や行動をコントロールする方法を身につける [24]．

Biofeedback や Neurofeedback は，ADHD の他，慢性疼痛，てんかん，高血圧など様々な症状において有効性が示されている [18]．

2.3 Decoded Neurofeedback

近年, Decoded Neurofeedback (DecNef) という手法が提案され [25], 様々な応用可能性が検討されている. DecNef の研究では, 機能的磁気共鳴画像法 (fMRI) により測定した fMRI 画像から脳の活動パターンを抽出し, 被験者に視覚的なフィードバック刺激を与え, 目標とする脳の活動パターンの生成法を被験者に学習させる. DecNef では, 機械学習を応用することにより fMRI 画像から脳の状態を推定し, 推定結果に基づいてフィードバックを生成する.

これまでに, 特定の脳機能を増幅するよう脳の活動を操作することが可能であることが示唆されている [26]. 機械学習を応用して生体情報から推定した脳の状態に基づいてフィードバックを生成する DecNef のような手法は, 様々な生体状態の操作への応用が期待できる.

2.4 Oculofeedback の提案

本研究では, 眼球運動をフィードバックする新たな Biofeedback のスキームとして Oculofeedback を提案する. Biofeedback では, 様々な生体状態を計測機器から推定しフィードバックすることにより, 被験者自身が自らの生体状態をコントロールする能力を訓練する. Oculofeedback では, 計測した眼球運動に基づいて注意などの脳の状態を推定し, 被験者が脳の状態を改善できるように訓練する. Biofeedback による多様な生体状態のコントロールや Neurofeedback や DecNef による脳の活動のコントロールの実現可能性が示唆されていることから, このような Oculofeedback による訓練には注意のような脳の機能をコントロールする効果が期待できる.

Oculofeedback は, Neurofeedback や DecNef と同様に, 脳機能を増強する訓練手法である. Oculofeedback の利点は, Neurofeedback や DecNef と

比べて容易に計測できる眼球運動から推定される脳の状態を利用して，注意や認知などの脳の高度な機能を訓練するための Biofeedback を実現できることである．

2.4.1 Oculofeedback システム

本研究では，まず，Oculofeedback を実現するシステムを開発する．一般的な Biofeedback 装置 [27] に基づいて図 2.1 のように Oculofeedback システムを構成する．

開発した Oculofeedback システムは，眼球運動計測装置とデータ処理装置，ディスプレイやスピーカーなどの刺激提示装置，ならびにそれらの装置を連携させる PC を上のプログラムによって構成されている．Oculofeedback システムを構成する装置には市販のものを使用する．各装置を連携させるプログラムは独自に作成したものをを用いる．このプログラムについては後述する．

Oculofeedback システムは眼球運動計測装置から得られる眼球運動データを実時間処理し，処理結果に基づき種々の感覚刺激を生成し，刺激提示装置に提示する．現在，様々な眼球運動計測装置や Web カメラを使用した眼球運動計測ソフトウェアが開発されていることから，本開発システムでは，特定の眼球運動計測装置に依存しないような高い汎用性を持つ仕様としている．開発したシステムは次の特徴を備えている．

- 特定のオペレーティングシステムに依存しないクロスプラットフォームで，Linux，Mac，Windows など様々な PC に対応可能
- 特定の眼球運動計測装置や刺激提示装置に依存せず，任意の端末をシステムに組み込む仕組みにより拡張可能
- 複雑な実験系もモジュール化された部品を組み合わせたスクリプトにより，簡単に柔軟な実験系を構築可能

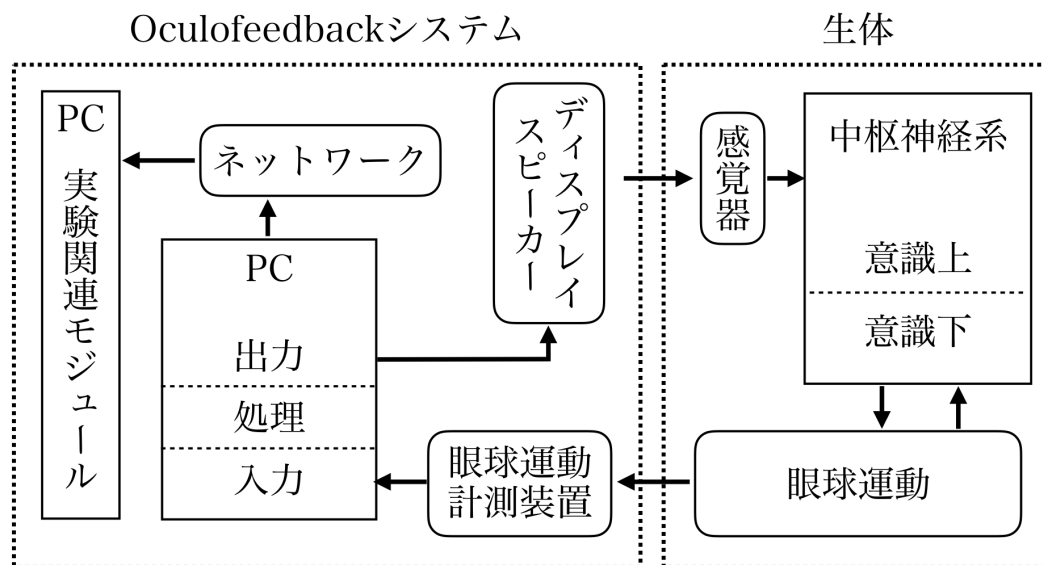


図 2.1: 開発した Oculofeedback システムの構成 ([27] の図を基に作成)

- プロセス間通信プロトコルを備え，異なる環境間での実行が必要な端末同士を組み合わせた実験系が構築可能

以下，使用した機器と開発したプログラムについて述べる．

2.4.2 対応機器

上述のように，本システムは様々な市販または自作の眼球運動計測装置を使用できるものとなっている．これまで，市販の EyeSeeCam (EyeSeeTec 社)[16] と Eye Tribe Tracker (The Eye Tribe 社) を用いて動作確認をしている．これらの他にも，眼球運動計測装置としては，眼球運動を計測中に，外部のプログラムがリアルタイムで眼球運動データを利用できる仕組みを備えていれば本システムに組み込むことができる．

後述するプログラムを実行し，計測した眼球運動データを処理したり，フィードバック刺激を生成したりするには，一般的な PC を使用する．

フィードバック刺激も様々な感覚刺激装置を介して与えられる．ここで

は、視覚刺激と聴覚刺激用の出力機能を実装した。一般的な PC とそれに接続したディスプレイとスピーカーを使って、眼球運動計測装置で計測された眼球運動データに基づく刺激の生成と提示ができる。

このように、開発した Oculofeedback システムでは汎用性と拡張性を考慮し、既存の機能に対して適切なプログラムを追加することにより、新たな入力端末や刺激提示装置に対応できる。

2.4.3 プログラム

本システムの核となるプログラムについて以下に述べる。

2.4.3.1 使用プログラミング言語

Oculofeedback を実現するには、リアルタイムで眼球運動を計測し、計測した眼球運動データを処理して刺激を生成し、フィードバック情報として被験者の感覚系に提示する必要がある。また、多様な実験系を構築するには、入力装置やデータへの処理内容、フィードバック刺激の変更などに、柔軟に対応できるプログラムが必要である。

これらの要件を満たすために、システムの実装には Python を使用した。Python には豊富なライブラリが備わっており、神経科学や心理学の分野などでも広く使われている。また、Python は、心理物理実験の実装用ライブラリである Vision Egg[28] や PsychoPy[29] などの開発にも利用されていることから、本システムの実装言語として最適であると考えられる。

データ処理や刺激の生成などのために使用した主なライブラリ名とその使用目的を表 2.1 に示す。

2.4.3.2 クロスプラットフォーム

使用する眼球運動計測装置や刺激提示装置によっては、特定のオペレーティングシステムが必要な場合がある。Python は Linux, macOS, Windows

表 2.1: 開発に使用した主なライブラリ

ライブラリ名	使用目的
Numpy	数値計算
Scipy	
VisPy	視覚刺激の生成
PyAudio	聴覚刺激の生成
msgpack	データのシリアルライズ
H5py	データの保存
PyQt4	GUI
gevent	疑似並列処理

それぞれに対応しており，使用したライブラリも主要なオペレーティングシステムでの実装が用意されていることから，開発したプログラムはクロスプラットフォームで利用できる．さらに，開発したプログラムは，各オペレーティングシステム上で動作するプログラムを TCP/IP 通信により接続することで，異なるオペレーティングシステム上のプログラムとの連携を可能にし，様々な装置を自由に組み合わせて実験環境を構築できる．

2.4.3.3 構造

本プログラムは，任意の入出力装置やデータ処理に対応する拡張性を実現するため，実験に必要な各機能をモジュール化した構造となっている．各モジュールは Python のクラスとして定義されており，必要に応じてインスタンス化することで機能を利用する．

柔軟な実験構成を実験者が簡単に実装可能とするには，リアルタイムのデータ入出力と処理を適切にモジュール化する仕組みが必要である．開発したプログラムのシステム設計には，実験に必要な各機能を提供するモジュールをアクターとするアクターモデルを採用した [30]．すなわち，各モジュール

ルは並列的に動作し、データのやり取りは各モジュール固有のキューを介して行われる。このような構造を採用することで、実験者は既存のモジュールに変更を加えることなく再利用でき、新たに必要な機能があれば容易に追加することが可能となる。

各モジュールの独立性を保つことで拡張が容易な構造となっているが、本プログラムの処理は必ずしも並列に動作する必要はない。そこで、並列化によるオーバーヘッドを回避するため、より効率的に動作すると考えられる疑似並列ライブラリ `gevent` によりアクターモデルを実装した。すなわち、各モジュールは独立したコルーチン上で疑似的に並列動作することとなる。以上のように、本プログラムはアクターモデルと疑似並列を採用し、拡張性と実行効率を実現する構造となっている。

以上の構造により、Oculofeedback システムは、図 2.1 に示したように、複数の PC 上に分散された入力、処理、出力モジュールや、その他実験に関連するモジュールがネットワークに接続され、各モジュールが協調して動作するようになっている。

2.4.3.4 使用方法

開発したプログラムは、Oculofeedback 実験の実施に必要な基本的な機能を API として提供するライブラリである。本プログラムを使用して実験を行う実験者は、プログラムのデータの入出力と処理を適切に設定するためのスクリプトファイルを Python により作成する必要がある。

コード 2.1 に開発したプログラムの API の使用例を示す。コード 2.1 では、眼球運動計測装置の EyeSeeCam から入力を受け取り、両眼水平眼球位置の平均を計算し、ストレージとディスプレイに計算結果を出力する例を示している。コードの 1, 2 行目では、開発したプログラムである `oculofb` というライブラリから、必要な機能を取り出している。EyeSeeCam は眼球運動計測装置から眼球運動データを受信するクラスであり、Storage と Display は実験記録やフィードバック刺激の出力先となるクラスである。Process ク

ラスには、実験者が実装する処理を補助する関数などが定義されている。4 から 12 行目では、この例の実験で使用することを想定した Gaze というデータ処理を定義している。Gaze クラスは、他のモジュールから入力されたデータに対して処理を行い、処理結果を次のモジュールへ送る。15 から 22 行目では各モジュールをインスタンス化し、インスタンスを組み合わせで適切なデータの流れを構成している。そして、24 行目以降ではデータの入力源となるインスタンスからメインループを開始している。このようなコードで、データをリアルタイムで扱う Oculofeedback 実験のためのプログラムが実装できる。少ないコードで簡単にデータのリアルタイム入出力と処理が実現でき、Oculofeedback 実験に必要な実験系を構成可能であることが分かる。

2.5 注意と眼球運動

本研究では、Oculofeedback の例として注意状態を操作する実験を実施する。ここでは、本研究で扱う注意について明確にするため、まず、注意の分類について述べる。また、眼球運動、特にマイクロサッカードと注意の関係について述べ、マイクロサッカードの発生特性の変調と注意の向上について説明する。

一般に注意という言葉は多様な場面で用いられるが、注意はその特徴に基づいて、いくつかの方法で分類することができる。本研究で扱う注意は空間的注意と呼ばれる注意のうち視覚と関わる注意である。

2.5.1 選択的注意 (Selective Attention)

選択的注意は多様な情報入力の中から特定の情報を強調する機能である。たとえば、写真の中から特定の人物を見つける際に、注意深く探す時の注意が選択的注意である。このとき、脳では対象人物を効率的に見つけるため

```

1 from oculofb import EyeSeeCam, Storage, Display
2 from oculofb import Process
3
4 class Gaze(Process):
5     def transfer(self):
6         data = self.receive()
7         # arbitrary data processing
8         data['gaze/hor'] = (
9             data['eyeseecam/left/hor'] +
10            data['eyeseecam/right/hor']
11        ) / 0.5
12        self.send(data)
13
14 if __name__ == '__main__':
15     esc = EyeSeeCam()
16     gaze = Gaze()
17     srg = Storage()
18     dsp = Display()
19
20     esc.connect(gaze)
21     gaze.connect(srg)
22     gaze.connect(dsp)
23
24     root = Root([esc])
25     root.start()

```

コード 2.1: Oculofeedback システム API の使用例

に、対象人物の特徴が選択的注意によって強調される．選択的注意によって、場所、物体、属性、大きさ、タイミングなど様々な特徴が強調されることで、注意対象に関する情報の処理が促進されると考えられている [31].

2.5.1.1 空間的注意 (Spatial Attention)

選択的注意のうち、空間における特定の位置に向けられる注意を空間的注意という [32]. ヒトの脳の情報処理では、視野の特定の位置における様々な特徴が並列的に処理されている．空間的注意はそのような情報の処理過程において、特定の位置にある特徴の処理を促進する機能であると考えられている．選択的注意のなかでも空間的な要素は特に強調されるものであると示されていることから [33]、空間的注意は注意の中でも特に重要なものであると考えられている [34].

2.5.1.2 視覚的注意 (Visual Attention)

選択的注意のうち、視覚と関わる注意を視覚的注意という．視覚的注意には空間的注意も含まれる．視覚的注意は視覚系の情報処理過程において、特定の位置や色、形の情報処理を促進する機能である．

近年の選択的注意に関する研究は、主に視覚的注意において活発に行われており、サルやヒトの視覚的注意の神経基盤について、神経活動の記録や脳波、脳画像を活用した神経科学の分野で調べられてきた [34].

2.5.2 トップダウン注意 (Top-down Attention) と ボトムアップ注意 (Bottom-up Attention)

注意が随意的なものであるか否かによって、トップダウン注意とボトムアップ注意に分類される．トップダウン注意は自らの意思で注意対象へ向けられる注意であり、ボトムアップ注意は与えられた刺激に反応して向けられる注意である．例えば自動車の運転中、交差点で左右を確認するときに向け

る注意がトップダウン注意で、突然飛び出してきた人に向けられる注意がボトムアップ注意である。トップダウン注意は経験や期待、現在の目標に基づく注意で、ボトムアップ注意は感覚刺激に反応する注意であると言える [35]。

ボトムアップ注意はトップダウン注意により抑制できることが示唆されている [36]。そして、トップダウン注意とボトムアップ注意は互いに独立した機構であり [37]、それらが互いに強く結びついている [38] と考えられている。よって、自らの意志で注意状態を変化できるトップダウン注意を変化させることで、ボトムアップ注意も制御できると考えられる。

2.5.3 顕在的注意 (Overt Attention) と 潜在的注意 (Covert Attention)

注意は、それに伴う随意的な眼球運動などの行動の有無に基づき、顕在的注意と潜在的注意に分類できる [39]。

注意対象へ注意が向けられた際に、その対象に視線を移動させた場合の注意は顕在的注意である。たとえば、自動車で角を曲がるとき、ほかの車両に衝突しないよう周囲を確認する場合には視線を向けて注意する。この時の注意は顕在的注意である。一方、視線を向けることなく注意だけを注意対象へ向けた場合の注意を潜在的注意という。たとえば、自動車の運転時に同乗者と会話するとき、運転者は視線を正面の道路に向けたまま同乗者に注意を向けることができる。この時の注意は潜在的注意である [40]。

2.5.4 注意の評価方法

注意を直接的に評価する方法は無い。そのため、注意は、注意が向いている状態と向いていない状態の差分としてその効果が評価される [41]。このような注意の効果を評価する方法として、Posner cueing paradigm[39] という

実験パラダイムが広く利用されている [32].

本研究で実施する実験で用いる課題 (第 4.3 項) を例に, Posner cueing paradigm について説明する. 図 2.2 は, Posner cueing paradigm における課題の流れの例である. 図の灰色の部分が被験者に提示する画面の背景であり, 矢印は画面の遷移を表す. この課題には, fixation, valid cue, invalid cue, reaction の 4 つの画面がある. 被験者はまず, 画面中央に表示された注視点 (十字の中心) を注視する. 数秒後, 画面に右または左を差す cue (矢先) が表示されるので, 被験者は cue の先へ注意を向けながらも, 視線は画面中央に維持し, 動かさない. ランダムな時間が経過したのち, 左右どちらかに表示されるターゲット (円) に視線を移す. 基本的な Posner cueing paradigm の流れは以上である.

このとき, ターゲットが表示されてから視線を移すまでの時間を反応時間 (RT) とする. 被験者は valid cue が表示された場合にターゲットに注意を向け, invalid cue が表示される場合にターゲットと逆に注意を向ける. 被験者の注意が適切に向けられていれば, valid cue の場合に RT が早くなり, invalid cue の場合に RT が遅くなることが知られている. Posner cueing paradigm の RT をみることで, 被験者が cue に適切に注意を向けられたかを評価できる.

上述した課題において, RT は空間的かつ視覚的な注意のパフォーマンスの評価指標であると言える. また, cue に基づいて被験者が向ける注意はトップダウンな潜在的注意であり, 表示されたターゲットに向ける注意は顕在的注意である.

2.5.5 注意に関わる眼球運動

本研究では, 主にサッカードとマイクロサッカードの 2 つの眼球運動を評価する. サッカードとマイクロサッカードは注意と密接に関係する眼球運動として知られている. ここでは, これらの眼球運動について述べる.

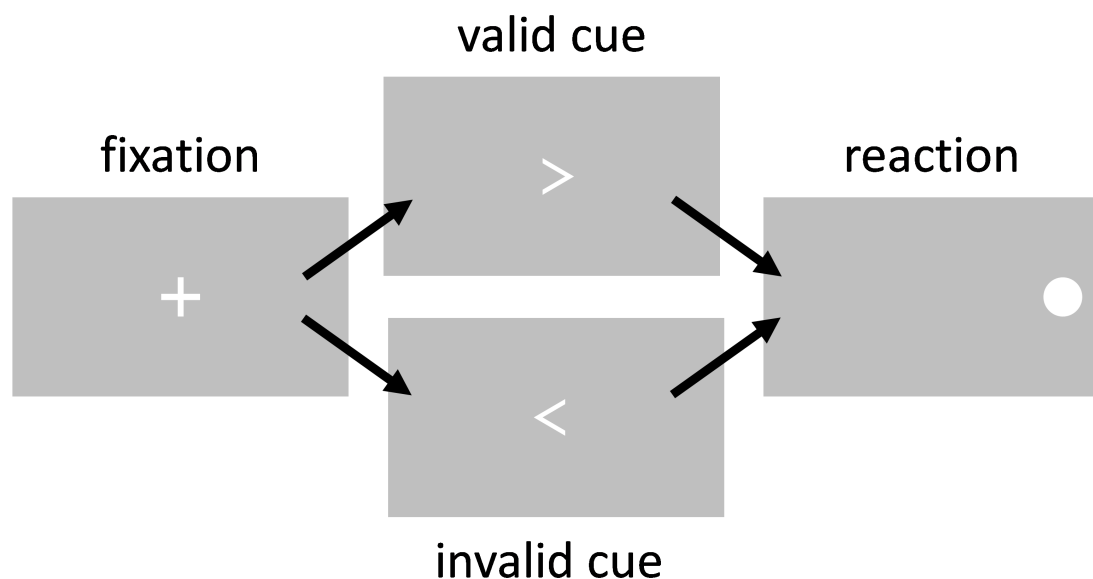


図 2.2: Posner cueing paradigm の流れ

2.5.5.1 サッカー

ヒトの視野は、鮮明に見える部分は限られていることから、何かを注意深く観察したい場合には、その対象に視線を移す必要がある。このときに発生する素早い眼球運動をサッカードという [42]。サッカードの生成に関わる脳の活動と潜在的注意に関わる脳の活動には、共通した活動がみられることが知られており [43]、サッカードの制御には注意が重要であると考えられている [44]。これまでの知見から、サッカードは注意と深く関連しており、互いに影響しあっていると考えられている。

2.5.5.2 マイクロサッカード

我々は通常、一点を注視していても、視線を完全に固定することはできず、常に小さな眼球運動が不随意に生じている [45]。このような注視中の微小眼球運動のうち、振幅 1 deg 以下の急速性のものがマイクロサッカード

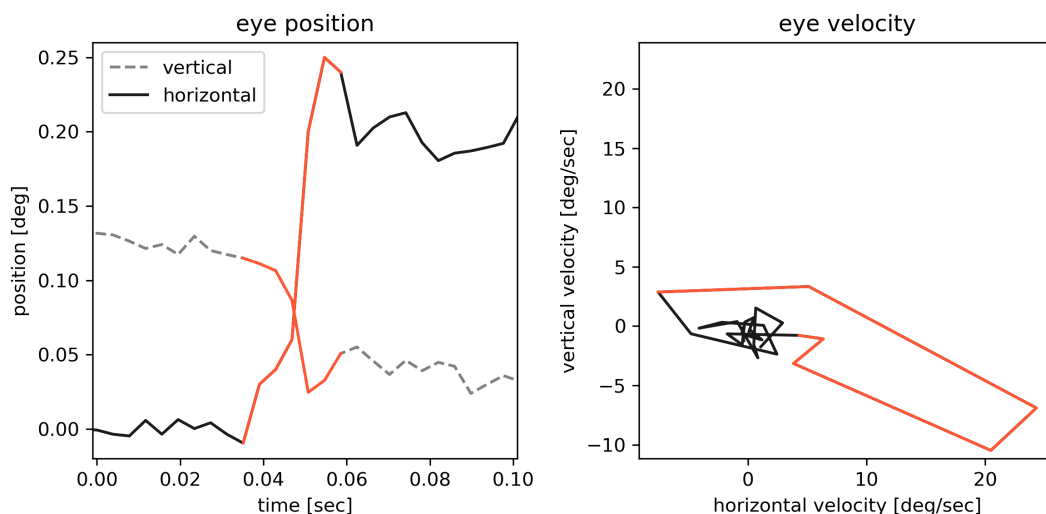


図 2.3: マイクロサッカードの例

(MSC) である [46]. MSC の生成機構はサッカードと共通していると考えられている [47] また, MSC の発生特性には, 潜在的注意の移動 [48] やその方向 [49] が反映されることが示されており, 潜在的注意と密接に関係していると考えられている [50, 51]. これらの知見は, 当初, 実験室内の統制された環境下で得られたものであったが, 近年, 自動車運転中のような, より自然な環境下でも MSC が発生し, その発生特性 (頻度, 方向, タイミング等) が歩行者の有無などの道路状況に依存して変化することが示されている [52, 53, 54, 55, 56]. したがって, MSC はヒトの潜在的注意の他覚的かつ定量的な指標として期待されている [57, 58].

眼球運動計測装置により計測した眼球運動データ中の MSC 波形の例を図 2.3 に示す. 図 2.3 左は MSC の水平 (horizontal) と垂直 (vertical) の眼球角度波形であり, 急峻に変化している橙色の部分 MSC である. MSC による眼球角度の変化量が MSC の振幅である. 図 2.3 右は左図と同じ MSC の眼球角速度波形であり, 注視中のデータ (黒色) に比べ MSC が急峻な変化を示していることが良く分かる.

2.5.6 マイクロサッカードに反映される注意

上述したように，サッカード，そしてマイクロサッカード (MSC) が注意と深く関連していると考えられている．潜在的注意と MSC を関連付ける研究結果から，以下のことが分かっている．

- Posner cueing paradigm の課題において，cue 表示後 300ms から 400ms における MSC の発生方向が cue 方向へ偏る [49]
- 同課題において，cue 表示後 200ms から 400ms の区間で MSC が発生した場合，RT が発生しなかった場合に比べて早い [58]
- 同課題において，ターゲット表示前に MSC が発生した場合，RT が遅くなる [59]

これらの知見に基づいて，適切に実験を設定することで，MSC の発生傾向から，課題遂行中の被験者の空間的注意状態の推定が可能であると考えられる．

2.6 まとめ

本章では，Oculofeedback の核となるプログラムについて説明した．開発したプログラムは Python で実装し，モジュール化された構造を採用することで汎用性と拡張性を備えつつ，実験者のプログラミング作業を効率化するシステム構成となっていることを示した．また，開発した Oculofeedback システムの使用方法を示し，Oculofeedback 実験の実装が容易に行えることを示した．そして，第 4 章で述べる Oculofeedback 実験で扱う注意と眼球運動について説明した．

第 3 章

Oculofeedback のための実時間 高精度マイクロサッカード 検出

本研究では Oculofeedback におけるフィードバックにマイクロサッカード (MSC) を用いる。本章では、まず、マイクロサッカードを高精度に検出する意義について述べる。次に、眼球運動の計測方法と計測されたデータから MSC を検出する難しさについて述べる。そして、従来法と提案法について解説し、その評価結果を示し、考察を述べる。

3.1 マイクロサッカードに基づく注意状態の推定

近年、眼球運動計測装置が比較的安価になったことから、心理学や神経科学の他、教育やスポーツ、自動車安全技術など、様々な分野で眼球運動が計測・評価されるようになってきている。なかでも、ヒトの注意状態推定が重要な分野において、眼球運動の実用化研究が進められている [5]。

第 2.5.5.1 項で述べたように、サッカードの発生には注意の移動が重要

[44]であることから、サッカードは注意の空間的移動を反映する指標と考えられる。こうした明確な視線移動を伴う注意（顕在的注意）に対し、視線移動を伴わない注意が潜在的注意である [40]。

従来、潜在的注意の評価には、サッカードのような容易に観測可能な評価指標がないことから、視覚刺激の変化に対する反応時間など、間接的な指標に頼った評価が行われてきた [48]。一方で、第 2.5.5.2 項で述べたように、MSC に潜在的注意が反映されていることが分かってきている。そのため、ヒトの潜在的注意の他覚的かつ定量的な評価指標として MSC を利用するために、MSC を高精度かつロバストに検出する手法が必要とされている。

潜在的注意のリアルタイムモニタや Oculofeedback など、MSC を様々な場面で応用するには、実時間で高精度の MSC 検出技術が不可欠である。しかしながら、MSC は微小かつ急速な眼球運動であり、自然な環境下で計測された眼球運動データから検出することは容易ではなく、これまで実時間での動作に適した、高精度で頑健な信頼性の高い MSC 検出法は存在していなかった。

3.2 眼球運動の計測とマイクロサッカード検出

現在、一般に普及している眼球運動計測装置は、カメラにより撮影した眼球映像から画像処理により眼球の視軸方向を眼位データとして出力する Video-Oculography (VOG) 方式によるものである [60]。VOG 方式の眼球運動計測装置では、赤外照明の空間的非一様性や環境光の非定常性、まぶたやまつげによる虹彩・瞳孔部分のオクルージョンなど様々な要因で眼位データにノイズが混入する。微小かつ急速で短時間に事象が終了する MSC を、こうしたノイズの混入したデータから高精度かつ実時間で検出することは挑戦的な課題であると考えられている [61, 62]。

これまで、いくつかの MSC 検出法が提案されてきたが、いずれも実験室内の高度に統制された環境で計測された眼位データからオフライン解析で検

出することを想定したものであり、人手によるパラメータ調節や計算コストの高い処理を必要とする [49, 61, 62, 63, 64].

本章の目的は、実時間 MSC 検出法を提案し、その有効性を示すことである。ここでは、まず、オフライン検出用に開発された従来法を実時間検出に応用する際の問題点について述べる。次に、それらの問題を解決するための畳み込みニューラルネットワーク (ConvNet) を用いた新しい実時間 MSC 検出法を提案する。ここでは、提案法で使用する ConvNet の構造とその学習方法について詳述し、さらに、実際の眼位データから構成したデータセットを用いて従来法と比較することにより、提案法の性能を評価し、その有効性を実証する。

3.3 従来法

MSC 検出法の中で、広く使われている手法は Engbert と Kliegl が提案した手法 (EK 法) である [49, 51, 63]. EK 法は、まず、眼位時系列データに MSC の特徴を強調するインパルス応答をもつフィルタを適用して特徴波形を算出する。そして、特徴波形に基づいて識別器が MSC の有無を判定し、眼位データの各時点において MSC の有無を表す 2 値のうちいずれかを出力する。EK 法では、異なるノイズレベルの眼位データに対応するために、特徴波形のノイズに対する感度を調節するパラメータ (λ) を人手により設定する必要がある。 λ の値を決定するための明確な基準はなく、経験的に $\lambda = 6$ がよく用いられている [62]. EK 法では、パラメータ λ に基づいて、適切な MSC 特徴量を算出することで、異なるレベルのノイズが混入した眼位データからの頑健な MSC 検出を実現している。

Otero-Millan らは MSC の特徴に対するクラスタリングを応用することで、人手によるパラメータの設定を必要としない手法 (OM 法) を提案している [61]. OM 法は、まず、眼位データから EK 法と同様の特徴波形を算出し、MSC 候補波形を抽出する。この候補波形から典型的な MSC 波形の特

徴に基づいて設計した特徴量を算出し、抽出した全候補波形の特徴量を入力としてクラスタリングを適用し、MSC 波形と非 MSC 波形を識別する。OM 法では、EK 法では考慮されていない MSC の特徴（発生頻度や瞬間最大速度、加速度など）を特徴量に組み込むことで、EK 法より高精度な MSC 検出を実現している。

最近では、Mihali らがベイズ推定を応用した MSC 検出法（BMD 法）を提案している [62]。BMD 法は、まず、典型的な MSC 波形の特徴に基づいて設計した統計モデルを用い、眼位データからベイズ推定によりモデルの状態を推定する。そして、モデルの状態に基づいて、各時点において MSC が発生している状態である確率を出力する。BMD 法では、MSC の特徴に加えて、ノイズの状態もモデルに組み込むことで、EK 法や OM 法に比べ、よりノイズの大きい眼位データからの MSC 検出が可能であることが示されており、確率的な出力値による頑健な MSC 検出が期待できる。

OM 法や BMD 法などの最新の MSC 検出法は、EK 法と比較して、典型的な MSC 波形の速度や加速度、MSC の発生頻度など、考慮する特徴を増やすことにより、高い MSC 検出精度を目指しているといえる。

実時間 MSC 検出法としては、Yuval-Greenberg ら [65] や Chen ら [66] の手法が挙げられる。しかしながら、Yuval-Greenberg らの手法は定期的に瞬きが含まれない 2 秒間の眼位データが必要なこと、Chen らの手法はノイズの変動に応じた実時間での人手による閾値の再設定が必要なことから、汎用的な MSC 検出法ではないと考えられる。

いずれの従来法も、頭部を固定して統制された環境下で計測される眼位データを想定して設計されており、非定常ノイズの混入を想定したものではないことから、より自然な環境下で計測される眼位データにおける有効性については検証が必要であると考えられる。また、OM 法と BMD 法はオフライン解析での利用を想定しているため、OM 法の場合は、クラスタリングの入力として全ての MSC 候補波形が必要であり、計算コストが高く、実時間検出への応用は困難である。BMD 法の場合は、眼位データの各時点におい

て確率的な収束を必要とする処理が実行され、実行時間が予測困難で計算コストも高いことから、これも実時間 MSC 検出には適さないと考えられる。

3.4 提案法

本研究では、実時間 MSC 検出における従来法の問題点を踏まえ、パラメータ調節が不要で、非定常ノイズを含む眼位データからも頑健に実時間で精度良く MSC を検出する方法を提案する。機械学習の分野では、識別対象の特徴を学習により獲得することで、高い精度の識別が可能となる例が数多く報告されている [67]。従来法では、典型的な MSC の特徴に基づいて特徴量が設計されたが、提案法では、眼位データから MSC の特徴を学習することで、MSC 検出精度の向上を図る。

MSC の特徴学習には、ディープラーニングを応用する。ディープラーニングは多様な分野で応用され、識別問題において高い精度を記録している機械学習手法であり [68]、高精度な MSC 検出の実現が期待できる。提案法は畳み込みニューラルネットワーク (ConvNet) を使用し、VOG 方式の眼球運動計測装置で計測された時系列眼位データから ConvNet のパラメータを自動的に学習することで、人手によるパラメータ調整を不要とする。また、提案法は OM 法と異なり、1 つの入力波形 (眼位データ) からでも出力 (MSC の発生有無) が得られるように設計できる。さらに、単純な畳み込み演算を適用することから、BMD 法に比べて高速な MSC 発生有無の判別処理が可能となる。

3.4.1 ConvNet の構造

提案法に用いる ConvNet の構造は、EK 法を参考に設計した。ConvNet は畳み込み層と畳み込み多層パーセプトロン層 [69] (MLPConv 層) の 2 つの部分により構成した。畳み込み層は EK 法におけるフィルタ処理

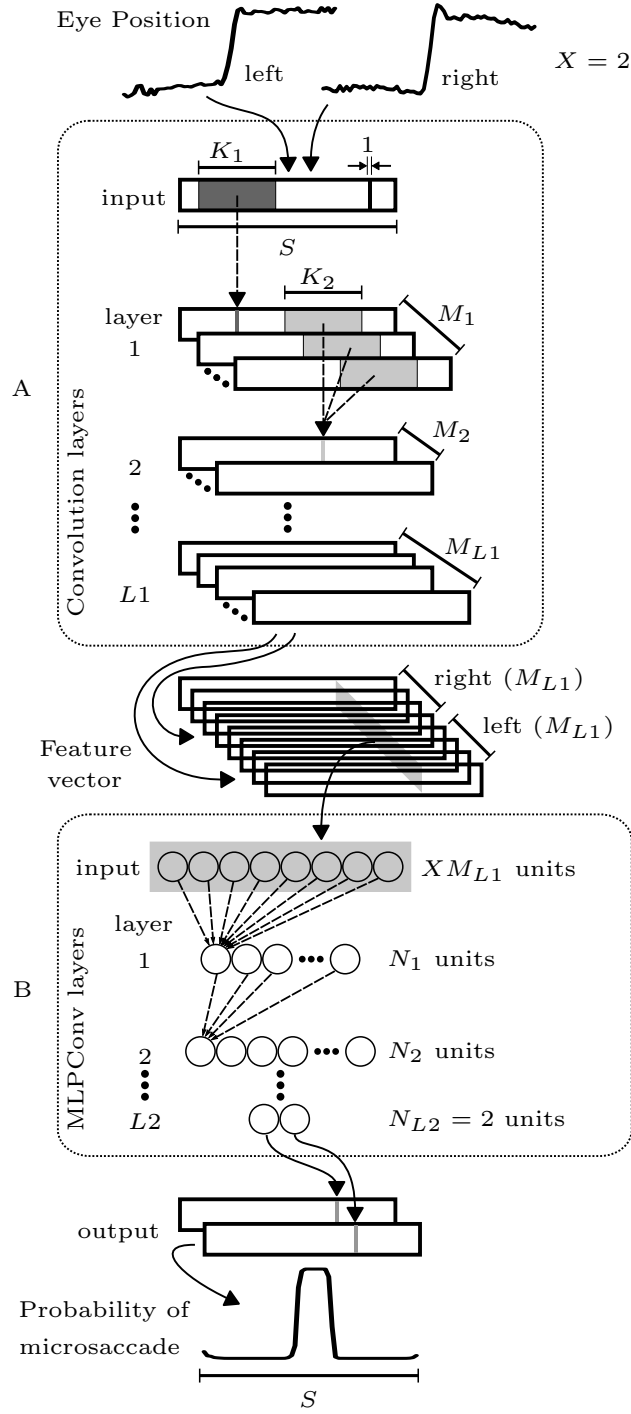


図 3.1: 提案法の畳み込みニューラルネットワークによる処理過程

に対応し，MLPConv 層は識別器に対応する．畳み込み層は，層の数 $L1$ ，フィルタ係数の個数 K_{l1} ，フィルタ個数 $M_{l1}(l1 = 1, 2, \dots, L1)$ で構成され，MLPConv 層は，層の数 $L2$ ，フィルタ係数の個数 1，フィルタ個数 $N_{l2}(l2 = 1, 2, \dots, L2)$ で構成される．活性化関数は，MLPConv 層の出力層が Softmax 関数，それ以外の層は ReLU 関数である．記号 $L1$ ， K_{l1} ， M_{l1} ， $L2$ ， N_{l2} が ConvNet の構造を決定するハイパーパラメータ，畳み込み層と MLPConv 層のフィルタ係数がパラメータであり，これらは第 3.4.2 節で述べる学習方法により最適化されるものである．

以下，畳み込み層と MLPConv 層で実行される処理について説明する．図 3.1 に，提案法で用いる ConvNet により，眼位データから MSC らしさが計算される過程を示す．ConvNet の入力は X 個の S 点で構成される眼位時系列データである（最上部 Eye Position）．出力は各時点における MSC らしさを表す確率（MSC 確率）で，入力波形点数と同様に S 点の波形（MSC 確率波形）となる（最下部 Probability of microsaccade）．図中の長方形は横軸を時間とした 1 次元の時系列データを表し，矢印はデータの流れを表す． X と S の値は，実験や応用で必要となる MSC 検出法の仕様に基づいて決定される．例えば，両眼のデータが使用可能で，水平方向の眼球運動のみを評価する場合は $X = 2$ とし，左右 2 つの水平眼位データを入力波形とする（図 3.1 の例）．

ConvNet は，最初に， X 個の入力波形にそれぞれ同一の畳み込み処理を適用し，特徴波形（Feature vector）を算出する（図中 A Convolution layers 部分）．図の場合，まず，右眼の眼位データを入力として 1 層目の K_1 点フィルタ M_1 個を適用し，特徴マップ数 M_1 の波形を得る．次に，その波形を入力として 2 層目の特徴マップ数 M_1 個の K_2 点フィルタ M_2 個を適用し，特徴マップ数 M_2 の波形を得る．同様にして， $L1$ 層目まで各層のフィルタを適用し，特徴マップ数 M_{L1} の波形を得る．そして，左眼の眼位データにも同一の処理を適用することで，最終的に特徴マップ数 XM_{L1} の特徴波形が得られる．この時，各層の入力波形の両端に 0 を $K_{l1}/2$ 点追加することで，

出力波形の点数を入力波形と同じ S 点としている.

畳み込み層から出力された特徴マップ数 XM_{L1} の特徴波形は, MLPConv 層 (図中 B MLPConv layers 部分) に入力される. このとき, MLPConv 層のフィルタ係数の個数は全て 1 とし, 畳み込み処理において時間的な依存関係を持たないようにした. ここで, 特徴マップ数 XM_{L1} の波形のある 1 点 (1 時刻) における MLPConv 層の全層の処理に注目すると, 入力特徴点数 XM_{L1} の多層パーセプトロン (MLP) と等価な処理となる. すなわち, MLPConv 層は, 特徴波形の各時点において, ユニット数 N_{l2} ($l2 = 1, 2, \dots, L2$) で構成される MLP を適用し, その時点における MSC 確率を出力することになる. 提案法では多値分類への拡張性なども考慮し, MSC 確率と非 MSC 確率を表す 2 つのユニットで構成した. 2 つのユニットの出力の和は 1 となることから, 両者のうち MSC 確率のみを ConvNet の出力として用いる.

ConvNet の入力波形数と識別対象の数は容易に拡張可能である. 例えば, 入力波形として両眼の水平および垂直方向の眼球運動を用いる場合は $X = 4$ となる. X を変更した場合, 畳み込み層の構造は同一で, MLPConv 層への入力特徴マップ数 XM_{L1} が変化する. また, MSC 以外に識別対象を増やし 3 値分類へ拡張した場合はユニット数が 3 となる. このように拡張した場合でも次節で述べる学習方法と同様の方法で, ConvNet を最適化することができる.

以上の構造により, 提案する MSC 検出法は, MSC の特徴を抽出するフィルタ (畳み込み層) と識別器 (MLPConv 層) を眼位データから学習する.

3.4.2 学習方法

提案法では ConvNet のパラメータ (畳み込み層および MLPConv 層のフィルタ係数) を, 眼位データを利用して学習し, ハイパーパラメータ ($L1$, K_{l1} , M_{l1} , $L2$, N_{l2}) を Random Search により探索する.

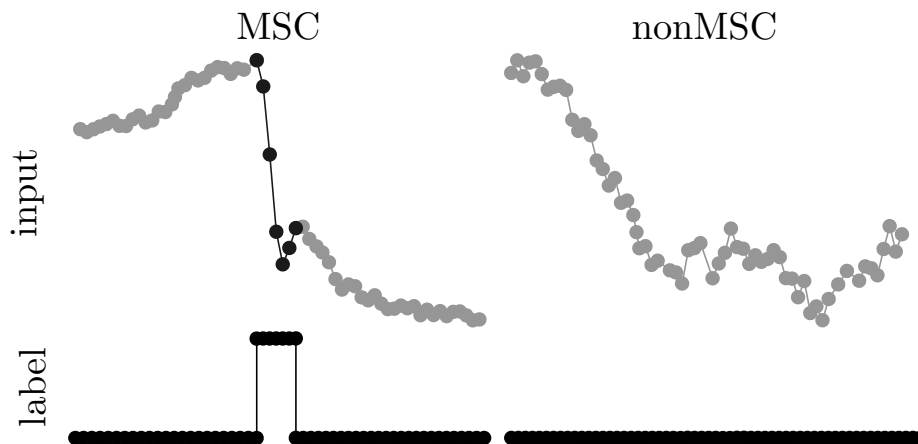


図 3.2: 学習に用いる入力波形とラベルの例

ConvNet のパラメータ最適化では，ニューラルネットワークの学習方法として推奨されている方法 [70] を参考に，バッチサイズ 50 のミニバッチを各イテレーションでランダムに生成して，ミニバッチ学習を実行する．最適化アルゴリズムには，確率的最急降下法的一种である Adam[71] を使用する．また，学習時には過学習を防ぐため Weight Decay（減衰率 wd ）を全てのパラメータに適用し，Dropout（Dropout 率 dr ）を入力層と畳み込み層の各層の出力に適用する．入力する眼位波形は波形ごとに，まばたきや計測エラー等による入力波形の欠損部を除いて平均 0，分散 1 になるように正規化する．正規化後，欠損部には 0 を代入する．教師データは入力波形の各時点における MSC の有無を表す 2 値波形（ラベル）である．

図 3.2 に入力波形とそのラベルの例を示す．左上が MSC を含む入力波形（MSC 波形），右上が MSC を含まない入力波形（非 MSC 波形）で，下がそれぞれのラベルである．入力波形の黒点部分は MSC が発生している箇所である．ラベルの立ち上がっている部分は MSC 発生箇所に対応する．入力波形のうち，MSC 発生中のデータ点を MSC 点（黒点部分），それ以外の点を非 MSC 点（灰色点部分）とする．ここで，MSC 波形中にも非 MSC 点が含まれるが，ConvNet においてフィルタを畳み込むことで，MSC 波形中の非

MSC 点には、MSC 点の情報も含まれることから、非 MSC 波形に含まれるものだけでなく、MSC 波形に含まれる非 MSC 点も学習に使用する。

入力波形の点数 S や MSC 発生箇所の点数は、使用する眼球運動計測機器のサンプリング周波数に依存するが、図のような入力波形を与えた場合、出力波形とラベルの全ての点の誤差を学習に使用すると、MSC 点に比べ非 MSC 点が多くを占め、パラメータの学習に大きな偏りが生じてしまう。そこで、 S 点で構成される入力波形とラベルから、ランダムに学習に用いる点を選択する。すなわち、MSC 点を p 点、MSC 波形における非 MSC 点を n_1 点、非 MSC 波形における非 MSC 点を n_2 点ランダムに選択する。これらの点数 $p, n_1, n_2 (> 0)$ をバランスよく設定し、学習に使用することで、MSC 点と非 MSC 点の学習の偏りを軽減する。学習時には S 点の入力波形を入力して算出した出力波形から、MSC 波形の出力の場合は MSC 点の p 点、非 MSC 点の n_1 点の誤差のみを、非 MSC 波形の出力の場合は非 MSC 点の n_2 点の誤差のみを学習に用いる。このとき、出力波形の各点には、畳み込み処理によりその周囲の情報が反映されていることから、選択された数点の誤差のみを学習に用いた場合でも MSC の特徴的な構造が学習できる。本研究における p, n_1, n_2 の具体的な設定法は 3.5.1.3 で述べる。

ハイパーパラメータは Random Search により探索する。表 3.1 に探索するハイパーパラメータと、その範囲を示す。探索の範囲は、何通りかのハイパーパラメータを学習した結果、学習の収束が確認できたハイパーパラメータ値の付近に設定した。ただし、出力層のユニット数 N_{L2} は 2 に固定である。

3.5 マイクロサッカード検出法の評価

現在、MSC 検出法の評価に必要なベンチマーク用データセットが存在しないことから、まず、MSC と非 MSC 波形にラベルを付加した MSC データセットを独自に作成した。次に、MSC データセットを使用して従来法の

表 3.1: ハイパーパラメータと探索範囲の一覧

ハイパーパラメータ		探索した値
畳み込み層の数	$L1$	1, 2, 3, 4, 5
フィルタ係数の個数	K_{l1}	5, 7, 9, 11, 13, 15
フィルタ個数	M_{l1}	4, 8, 16, 32, 64
MLPConv 層の数	$L2$	2, 3, 4
ユニット数	N_{l2}	4, 8, 16, 32, 64, 128
減衰率	wd	0 から 0.01
Dropout 率	dr	0 から 0.1

検出精度を評価した．さらに，提案法の ConvNet を MSC データセットにより学習させ，検出精度を評価して従来法と比較した．以下，詳細を記す．

3.5.1 マイクロサッカードデータセット

ベンチマーク用に MSC データセットを以下のように作成した．データセットは 6 名の被験者（20 代男性，A, B, C, D, E, F）から計測した眼位データ，眼位データの各時刻における MSC 発生有無のラベル，ConvNet の学習と MSC 検出結果の評価対象となる MSC 波形と非 MSC 波形から構成される．

3.5.1.1 眼位データ

眼位データは，VOG 方式のアイトラッカーである EyeSeeCam（EyeSeeTec 社）[16] を使用して計測した．EyeSeeCam の眼球角度の解像度は最小 0.01 deg であり，MSC の計測に十分な精度を有する．計測した眼位データは，被験者に水平方向の潜在的注意を誘発する Posner cueing

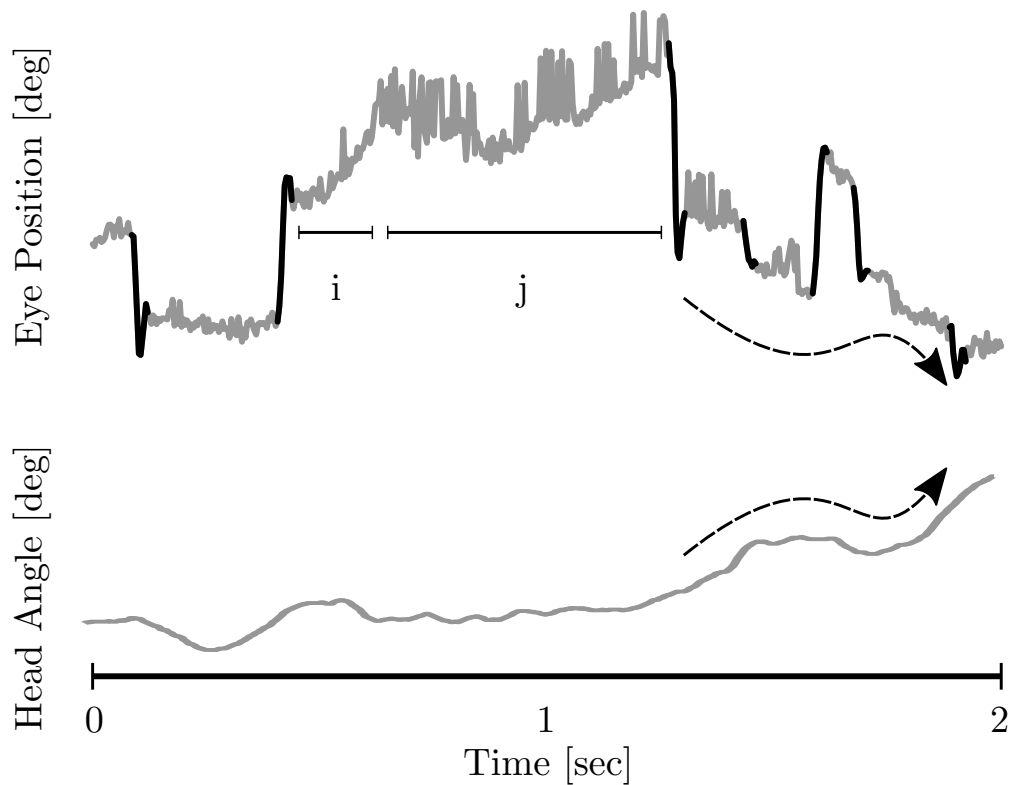


図 3.3: 水平眼位と頭部回転計測データ例

paradigm[39] の課題を与えながら，正面に提示した注視点を注視してもらい，サンプリング周波数 256 Hz で，両眼の水平および垂直眼位を約 20 分間記録したものである．このとき，被験者は椅子に座り，専用の台に顎を乗せた状態で楽な姿勢をとり，頭部は固定しなかった．これにより，計測された眼位データには被験者の姿勢変化や頭部動揺などに起因したノイズが混入した．また，被験者 D はまぶたやまつげの状態により計測装置との相性が悪く，他の被験者に比べてノイズ混入の頻度が高い眼位データとなった．

図 3.3 に計測した眼位データの例（被験者 A）を示す．黒線部分が MSC の発生箇所を表す．図から，頭部の動きによる眼位の変動（矢印）やノイズレベルの変動（区間 i, j）などが確認できる．こうした MSC データセットを使用することにより，頭部固定などの強度の統制条件下で計測されたデー

タと比べ、より自然な眼球運動時に発生する MSC の検出評価が可能となると考えられる。

3.5.1.2 ラベル

MSC 発生有無のラベルは、被験者の両眼の眼球角度波形を PC スクリーン上に提示して目視で確認し、0.1 deg 以上の変化が 4 点 (15.625 ms) 以上、両眼で同時に同方向へ急速に眼球角度が変化している箇所を MSC 発生箇所とした [72]。このとき、必要に応じて、同時計測された両眼眼球映像も参照し、確実に MSC が発生していることを確認した。そして、波形を構成する各点に、MSC 発生箇所では 1、それ以外では 0 とするラベルを付与した。

3.5.1.3 入力波形

MSC 波形と非 MSC 波形は、提案 ConvNet の入力波形点数と同じ S 点の波形である。本研究では、今回の測定条件において MSC 1 つ分の特徴が捉えられる波形の長さとして $S = 64$ (0.25 秒) とした。

MSC 波形と非 MSC 波形は、計測した眼位データから以下の手順で被験者別に切り出した。MSC 波形は、計測した眼位データ中の全ての MSC について、MSC 発生箇所が中央に配置されるように S 点を切り出して作成した。非 MSC 波形は、MSC 発生箇所が含まれず、かつ欠損部が S 点の 30 % 以上とならない波形を、重ならないようランダムに S 点を切り出して作成した。6 名の被験者から得られた MSC 波形と非 MSC 波形の例を図 3.4 に示す。

3.4.2 節で述べたように、MSC 点と非 MSC 点の学習の偏りを軽減するために、学習に用いる点数 p , n_1 , n_2 をバランスよく設定する必要がある。データセットに含まれる MSC のうち、MSC 発生箇所の点数が最小なものは 3 点 (約 10 ms) であったことから、MSC 波形からランダムに選択する MSC 点の数は $p = 3$ とした。

ConvNet のパラメータがミニバッチ学習により最適化されているこ

とから，MSC 波形と非 MSC 波形における学習に用いる点数が等しい ($p + n_1 = n_2$) と，1 回の学習で実行される処理が固定され，プログラムの実装と計算効率などの面で有利である．一方，MSC データセットに含まれる MSC 波形の数（発生した MSC の数）を a ，非 MSC 波形の数を b とすると，MSC 点数 (p) と非 MSC 点数 (n_1, n_2) のバランスが良い比率は $1 : 1 = ap : an_1 + bn_2$ である．このうち， a, b, p は定数であるので， $n_1 < p$ であり， n_1 は 1 もしくは 2 となる． $n_1 = 1$ とした場合，比率の式を変形して $ap = an_1 + bn_2$ となることから， $(3-1)a = 4b$ で， $b = a/2$ となる． $n_1 = 2$ とした場合， $(3-2)a = 5b$ から， $b = a/5$ となる．ConvNet の学習において，入力波形数はできるだけ多い方が良いため， $n_1 = 1$ ， $n_2 = p + n_1 = 4$ とし，非 MSC 波形の数は $b = a/2$ とした．

6 名の被験者の眼位データから抽出した入力波形の数を表 3.2 に示す．MSC 波形が 4259 個，非 MSC 波形が 2126 個である（計 6385）．非 MSC 波形数が正確に MSC 波形数の半分になっていないのは，計測した眼位データに MSC の発生数や欠損部が多く，非 MSC 波形を切り出せる区間が不足したためである．

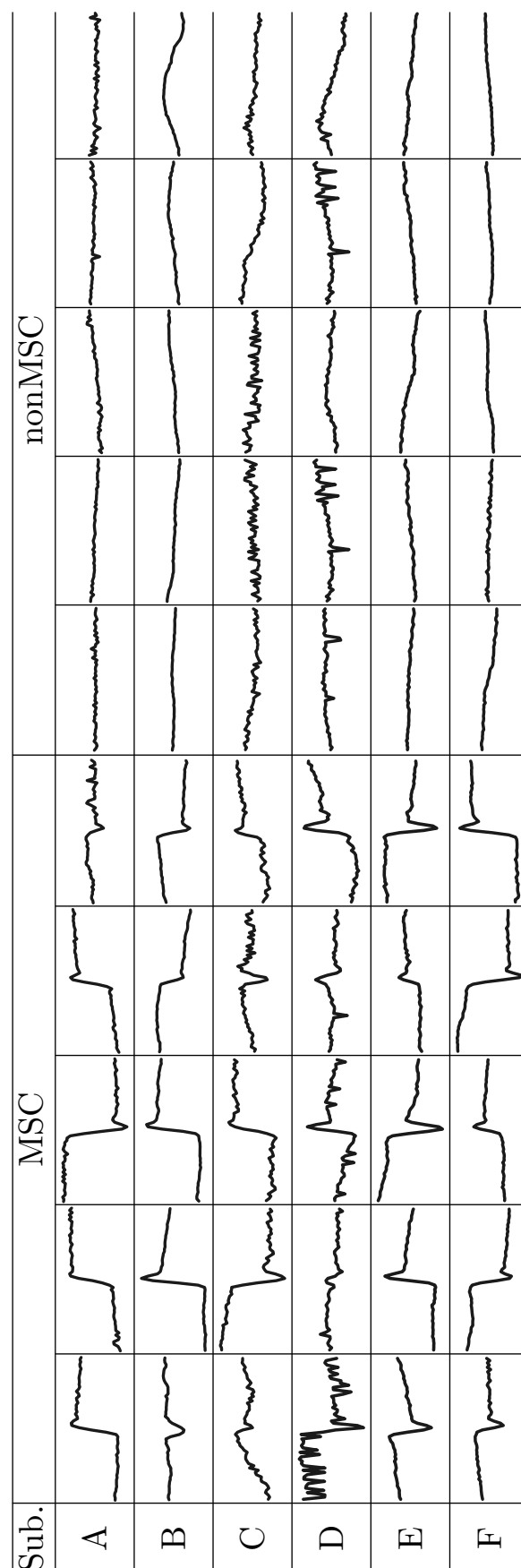


図 3.4: MSC 波形と非 MSC 波形の例

表 3.2: 学習と検出精度の評価に用いる入力波形数

被験者	MSC 波形	非 MSC 波形
A	1440	720
B	281	140
C	891	445
D	456	227
E	581	290
F	610	304
計	4259	2126

3.5.2 検出精度の定義

MSC 検出法の精度は，MSC データセットの各波形に対する MSC 検出結果に基づいて，以下のように定義する（図 3.5）．すなわち，MSC の検出結果について，MSC 波形では，1 点でもラベルと同じ点で MSC と判定すれば正解とし，本来 MSC と判定すべき箇所でも 1 点も MSC と判定できなければ不正解とする．非 MSC 波形では，1 点も MSC と判定しなければ正解とし，1 点でも MSC と判定すれば不正解とする．全波形のうち正解した波形の割合を検出精度とする（図中の例で検出精度を算出した場合は 50 % となる）．

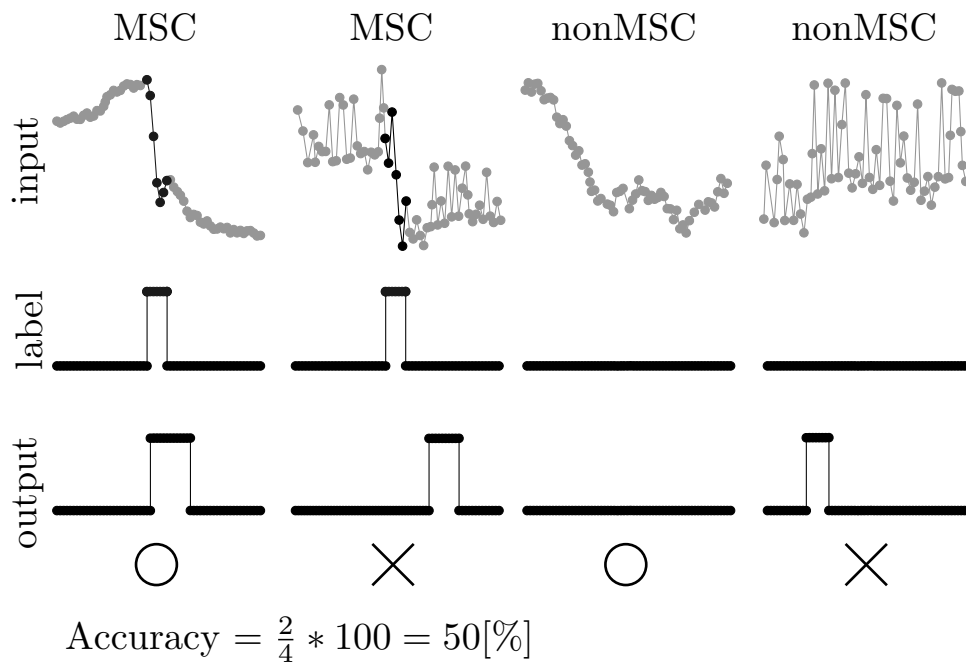


図 3.5: MSC 検出精度の計算方法

3.5.3 従来法の評価

3.5.3.1 方法

従来法のうち、OM 法と BMD 法、ならびに EK 法の検出精度について、作成したデータセットを用いて評価した。

OM 法は <http://smc.neuralcorrelate.com/sw/microsaccade-detection> で公開されている実装を利用し、MSC データセットの両眼の水平・垂直眼位データを入力して MSC を検出した。

BMD 法は <http://github.com/basvanopheusden/BMD> で公開されている実装を利用し、両眼の水平眼位データを入力して MSC を検出した。

EK 法は、Python の numpy パッケージを利用して実装し、両眼の水平眼位データの波形を入力して MSC を検出した。EK 法では手動によるパラメータ入の設定が必要であることから、以下の 2 つの方法で入を設定し

表 3.3: 従来法の MSC 検出精度

被験者	OM 法 [%]	BMD 法 [%]	EK6 法 [%]	EKopt 法 (λ) [%]
A	51.6	55.9	94.0	94.1 (6.3)
B	43.1	91.4	98.3	98.3 (5.4)
C	42.9	70.1	93.7	94.5 (6.4)
D	43.0	64.3	86.5	87.7 (5.4)
E	63.5	81.7	98.4	99.9 (8.1)
F	78.9	83.5	98.9	99.3 (6.3)

た．一つは， λ の値をよく使われる 6 と設定し，EK6 法とした．もう一つは， λ を 1 から 16 まで 0.1 刻みで変化させて算出した検出精度に基づいて，最適な λ を設定し，EKopt 法とした．すなわち，MSC 検出精度は EK6 法 $\leq$ EKopt 法となる．

3.5.3.2 結果

従来法を各被験者のデータに適用し検出精度を求めた結果を表 3.3 に示す．EKopt 法については括弧内に設定した最適な λ の値を示した．表 3.3 から，OM 法と BMD 法は，EK 法と比較してどの被験者のデータに対しても検出精度が低いことがわかる．OM 法では，眼位データに混入したノイズによって MSC の頻度や加速度などの特徴量の算出が困難となり，検出精度が低くなったと考えられる．Mihali らによる OM 法の評価でも，ノイズの混入により OM 法の検出精度が低下することが示されている [62]．BMD 法では，正規分布に従うノイズの混入を考慮した統計モデルにより，頑健な MSC 検出を実現したが，今回のデータセットのように非定常非正規ノイズが混入するデータでは MSC 発生状態の推定が困難となったことで，検出精

度が低くなったと考えられる．また，これらの手法は頭部を固定した状態における眼位データを想定して特徴量やモデルを設計していることから，頭部を固定しないことにより発生する図 3.3 のような眼位の変動（矢印）により検出精度が低下したと考えられる．

一方，EK 法では全体としておよそ 90 % 以上の検出精度が得られている．EK 法では，ノイズに影響されやすい瞬間最大速度や加速度などの特徴を用いず，他の従来法と比較して頑健な MSC 特徴量のみを利用したことにより，ノイズの混入や頭部動揺に伴う眼位変動が含まれる眼位データに対しても，他の手法より高精度な検出が可能となったと考えられる．しかしながら，EKopt 法では最適な λ を採用したにもかかわらず，EK6 法と比べて検出精度があまり改善しなかったことから，EK 法の検出精度には限界があると考えられる．特に，機器との相性が悪く，大きくノイズが混入した被験者 D のデータセットで検出精度が低いことから，ノイズの混入が激しい場合には，EK 法による高精度の MSC 検出は困難であると言える．

3.5.4 提案法の評価

3.5.4.1 方法

従来法の評価結果から，提案法を EK 法と比較評価する．ConvNet の学習では，一般的にデータセットを，パラメータの最適化に使用する学習用データ，学習の収束を検証するための検証用データ，ならびにパラメータやハイパーパラメータの決定に用いない試験用データに分割する．ConvNet の識別性能は試験用データで評価される．

学習・検証・試験用データの分割方法は，学習した ConvNet の検出精度に影響する．ここでは，提案法が EK 法と比較して頑健な手法であることを示すため，以下のように被験者別に分割した．まず，EK 法については，EK6 法と EKopt 法の 2 つの場合について評価したが，パラメータの設定にデータが関与しない EK6 法を試験用データとの比較対象とする．そこで，

EK6 法で検出精度が最低となった被験者 D と最高となった被験者 F のデータを試験用データとした．このような試験用データにより，EK 法による最高の検出精度と提案法の検出精度を比較すると同時に，EK 法による MSC 検出が困難である場合の提案法の検出精度を評価する．

次に，検証用データに波形数をもっとも多い被験者 A のデータを使い，学習用データにそれ以外のデータ（被験者 B, C, E）を使うこととした．このように分割することにより，できるだけ多くの被験者のデータを利用することで汎化性能を高めながら，できるだけ多くの波形で ConvNet の精度を検証することが可能となる．

上述したデータ分割方法により，データセット全体で 4259 個の MSC 波形と 2126 個の非 MSC 波形は，学習用データの MSC 波形数が 1753，非 MSC 波形数が 875，検証用データの MSC 波形数が 1440，非 MSC 波形数が 720，試験用データの MSC 波形数が 1066，非 MSC 波形数が 531 に，MSC 波形と非 MSC 波形の比率を保ったまま分割される．また，本評価方法により分割された学習・検証・試験用データには，それぞれ異なる被験者のデータが含まれ，検証・評価データでは学習に使用したデータと異なる被験者のデータが使用されることから，データを取得した人物に依存しない MSC 検出精度の評価が可能である．

本研究では，MSC データセットの作成のために計測した眼位データが，水平方向の潜在的注意を誘発したものであり，主要な MSC の発生方向が水平方向であることから，両眼の水平眼位データを学習に用いた．したがって，ConvNet の入力波形数 X は 2 である．入力波形は，3.5.1.3 項で述べた MSC 波形と非 MSC 波形であり，教師データはそれらの波形に対応するラベルである（図 3.2，図 3.5. $S = 64$ ）．

提案 ConvNet は TensorFlow で実装し，3.4.2 節で述べた方法により，パラメータとハイパーパラメータを学習した．提案 ConvNet の学習は CPU (Intel Core i7-4790K) と 2 つの GPU (NVIDIA GeForce GTX TITAN X) を備える PC により実行した．Random Search によるハイパーパラメータ

表 3.4: 提案法と従来法（EK 法）の検出精度

データ	被験者	提案法 [%]	EK 法 [%]
学習	B	99.0	98.3
	C	99.8	94.5
	E	100.0	99.9
検証	A	96.6	94.1
試験	D	94.6	86.5
	F	98.9	98.9

の探索は 2 個の GPU 上で並列に実行した．各ハイパーパラメータの組み合わせにおいて，学習用データを用いて ConvNet のパラメータを最適化し，学習が収束した時点での検証用データにおける MSC 検出精度を，その組み合わせによる検出器の検出精度とした．

3.5.4.2 結果

2 日間にわたり，363 通りのハイパーパラメータで学習させた結果，最大の検出精度が得られたハイパーパラメータは，畳み込み層数 L_1 が 5 層で，1 層目から順にフィルタ係数の個数 K_{l_1} が 15, 13, 13, 9, 11, フィルタ個数 M_{l_1} が 32, 8, 64, 64, 16, MLPConv 層数 L_2 が 3 層で，1 層目から順にユニット数 L_{l_2} が 16, 128, 2 であった．また，減衰率 wd は 0.000032, Dropout 率 dr は 0.09 であった．

表 3.4 に提案法と EK 法の検出精度を学習・検証・試験用データ別に示す．EK 法の学習・検証用データの欄には，パラメータの最適化にデータが関与している EKopt 法の検出精度を示した．試験用データの欄には，パラメータをデータに関わらず設定した EK6 法の検出精度を示した．

表 3.4 より，提案法は全ての学習用データにおいて 99 % 以上の検出精度

を示していることから、提案 ConvNet は各被験者の多様な MSC (図 3.4) の特徴を学習可能な構造であると考えられる。また、検証用データにおいては、EK 法に比べて高い精度が得られていることから、学習により MSC 検出に有用な特徴が獲得されたと考えられる。試験用データにおいては、計測データへのノイズ混入が激しかった被験者 D で、提案法が EK 法と比べて、検出精度が 8.1 % 向上したことから、提案法は従来法より高精度な MSC 検出法であると言える。さらに、EK 法では大きく異なる検出精度を示した被験者 D と F を含め、提案法では学習・検証・試験用データの全てにおいて検出精度が安定して高いことから、本提案法は、様々な MSC の波形プロファイルを持つ異なる被験者の眼位データ [73] に対しても、より優れた汎化性能を有するものと考えられる。

提案法による MSC 検出速度を、実時間 MSC 検出を想定し、EyeSeeCam による眼球運動の計測が可能な MacBook Air (CPU Core i7 1.7 GHz) 上で評価した。MSC の検出に要する処理時間は Python の `timeit` モジュールを利用して計測した。その結果、64 点の波形に提案法を適用した場合、MSC 確率波形の出力に必要な処理時間は 1.98 ms であった。サンプリング周波数が 256 Hz の場合、新たにデータが 1 点取得できるまでの時間は $1/256 \text{ s} = 3.9 \text{ ms}$ である。提案法の処理時間はこれより短いことから、提案法の処理速度は実時間 MSC 検出法としても十分であると考えられる。

EK 法についても同様の環境で処理時間を計測した結果、0.18 ms であった。OM 法と BMD 法の処理時間は、これらの手法が 64 点のようなデータ点数の少ない波形の入力に対応していないことから、計測不能とした。

3.6 考察

EK 法のフィルタと ConvNet の学習により獲得された畳み込み層一層目のフィルタのインパルス応答とその周波数特性を各々図 3.6, 3.7 に示す。これらの図から、ConvNet には様々な特性をもつフィルタが獲得されたこと

がわかる．また，図 3.6, 3.7 中の各々星印のついたインパルス応答と周波数特性（1 行 2, 3 列目）を見ると，EK 法で使用されるフィルタ（各図最右列）とよく似た特性が獲得されているように見える．提案法では，複数の被験者から得られた実データにより構成したデータセットを用いて ConvNet を学習することで，従来法に比べより実データに適したフィルタ特性を獲得したものと考えられる．

ニューラルネットワークを用いた検出器は一般的に，データセットを更新することによってパフォーマンスを改善できることから [70]，提案法も同様に，より多くの被験者から多様な環境下における眼球運動を計測してデータセットを更新することにより，様々な使用状況において頑健に MSC を検出可能なものに改善可能であると考えられる．

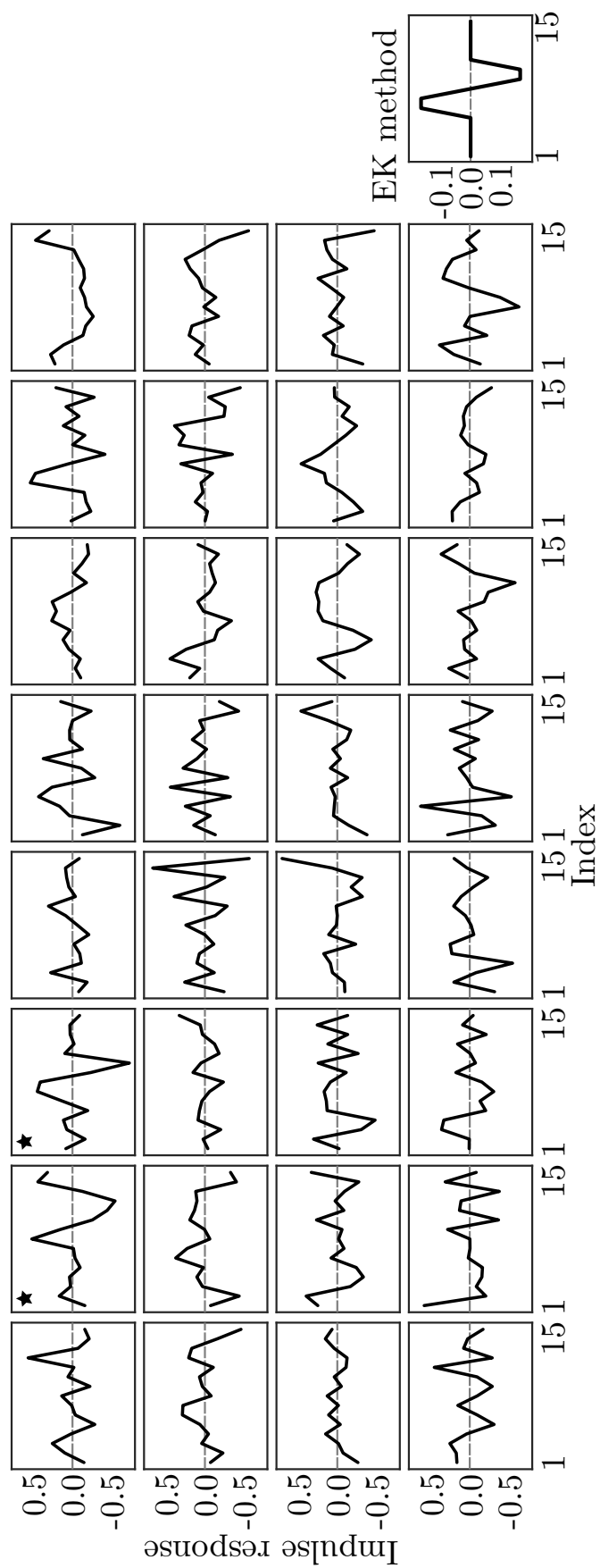


図 3.6: EK 法のフィルタと学習した ConvNet の一層目のフィルタのインパルス応答 ($K_1 = 15$)

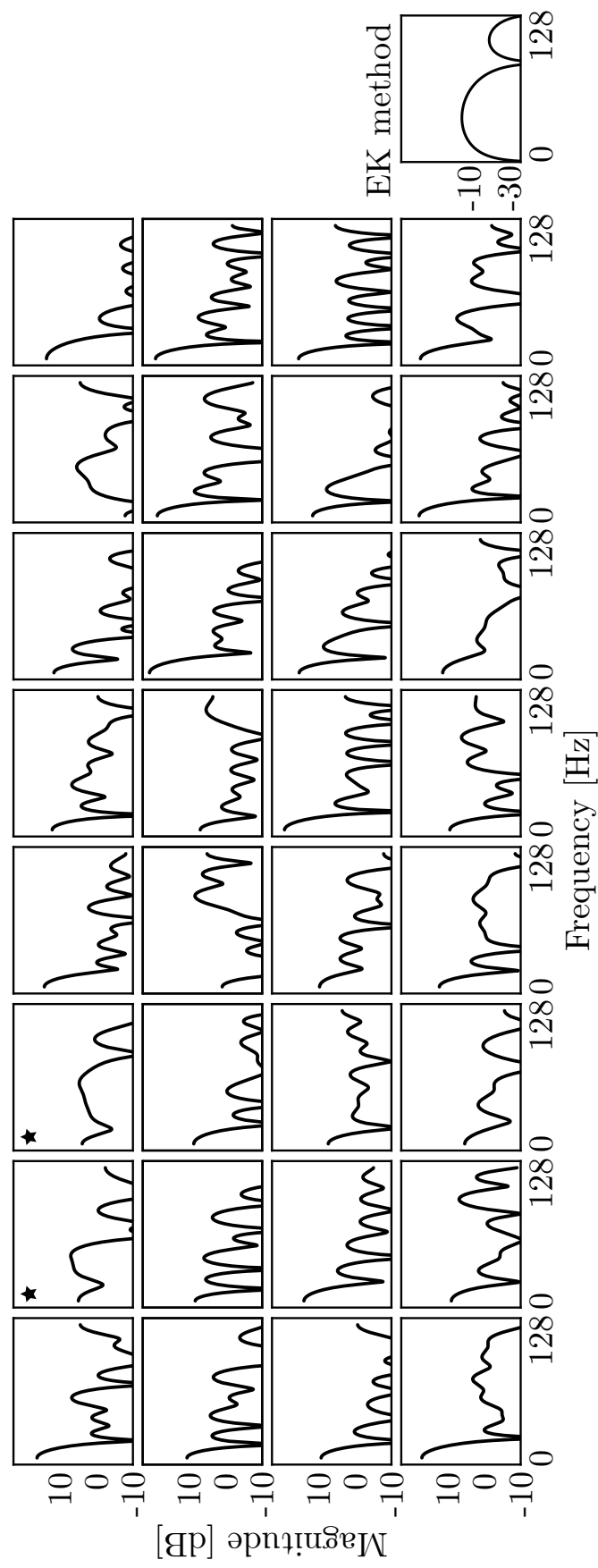


図 3.7: EK 法のフィルタと学習した ConvNet の一層目のフィルタの振幅特性 (サンプリング周波数 : 256 Hz)

第 4 章

Oculofeedback による注意状態 操作実験

開発した Oculofeedback システムを使用して注意力を訓練する Oculofeedback 実験を実施した．本章では，実施した Oculofeedback 実験の目的，方法，結果について述べ，考察する．

4.1 マイクロサッカードフィードバック実験

本研究では，Oculofeedback 実験として，マイクロサッカード (MSC) から推定した注意状態に基づいてフィードバックを与える，マイクロサッカードフィードバック実験を実施した．

4.1.1 Oculofeedback による注意操作の仕組み

Oculofeedback において，マイクロサッカードに基づいて被験者の注意力を向上させる訓練は以下の手順に従う．

1. 被験者に注意のコントロールが必要な課題を与える

2. 課題遂行中の被験者から眼球運動を計測する
3. 計測した眼球運動からマイクロサッカードを検出する
4. 検出したマイクロサッカードの発生特性に基づき，被験者の注意状態を推定する
5. 推定した注意状態を反映する刺激を被験者に提示する
6. 被験者は課題中の自身の注意状態を把握し，より良い注意状態に近づくように務める
7. 以上の 1 から 6 を繰り返す

上記手順の 5 がフィードバックに当たり，その刺激の提示は実験・訓練目的に応じて，課題中逐次的に，もしくは課題終了時に行う．被験者は上記のように繰り返す課題の中で，どのような意識や行動が望ましい注意状態に結びついたかを学習する．そして，学習を続けることにより，被験者自身が注意をコントロールする能力を獲得することが期待される．以下，この手順に基づいた実験の目的と内容を述べる．

4.2 目的

本実験の目的は，以下に述べる Oculofeedback 実験の課題における被験者の Reaction Time (RT) と課題遂行中の MSC の発生特性の関係を解析することで，Oculofeedback による注意力向上訓練の実現可能性を検証することである．

本実験では，与えられた課題において，被験者の注意状態を最適な状態へ誘導するために，被験者に MSC の発生特性に基づく注意状態をフィードバックする．訓練に繰り返し取り組むことにより，被験者の課題における RT が早くなることを期待している．

4.3 方法

被験者は Posner cueing paradigm に基づいた課題を，1 回の実験で一定の条件を満たすまで繰り返し実施する．

図 4.1 に課題の流れと画面の表示例を示す．画面に示した指標を，左から順に Fixation (十字)，Cue (矢印の先)，Target と呼ぶ．各指標の大きさは，Fixation の視野角が 0.73° ，Cue の縦横が $0.98^\circ \times 0.61^\circ$ ，Target の視野角が 0.73° である．表示する Target の方向を指す矢印を Valid Cue，逆の矢印を Invalid Cue とする．Fixation と Cue は画面中央に表示し，Target は中央から 12.4° 離れた左右どちらかに表示する．課題の画面は次のように切り替える．まず，Fixation ターゲットを 1.5 から 2.5 秒の間のランダムな時間表示する．その後，Cue を 2.0 から 2.5 秒の間で同様にランダムな時間表示する．そして，Target を 1 秒間表示する．Target の表示方向は課題開始時にランダムに決定し，Valid Cue を 8 割，Invalid Cue を 2 割の確率で表示するように設定する．

被験者の注意状態のフィードバックとして，Target 表示後に次のような刺激を与える．図 4.2 は，Target の表示終了後に与えるフィードバック画面の説明図である．画面には注視点である十字と同心の二つの白い円が表示される．黒い矢印は，内側の円が課題を重ねるごとに大きくなる様子を表している．フィードバックは，この内側の円が大きくなり，スピーカーから音が鳴る場合と，円が変化せずスピーカーから音も鳴らない場合の 2 パターンの視聴覚刺激である．

フィードバックは，被験者が Fixation と Cue を注視している区間，Target が表示される前までの眼球運動から検出した MSC に基づいて決定される．Cue が表示されてから t_1 秒以上 t_2 秒未満の間に MSC が発生していた場合，スピーカーから音を鳴らすと同時にディスプレイ上の同心円の内側の円を大きくする．一方，この間に MSC が発生しなかった場合，音は鳴らさず，

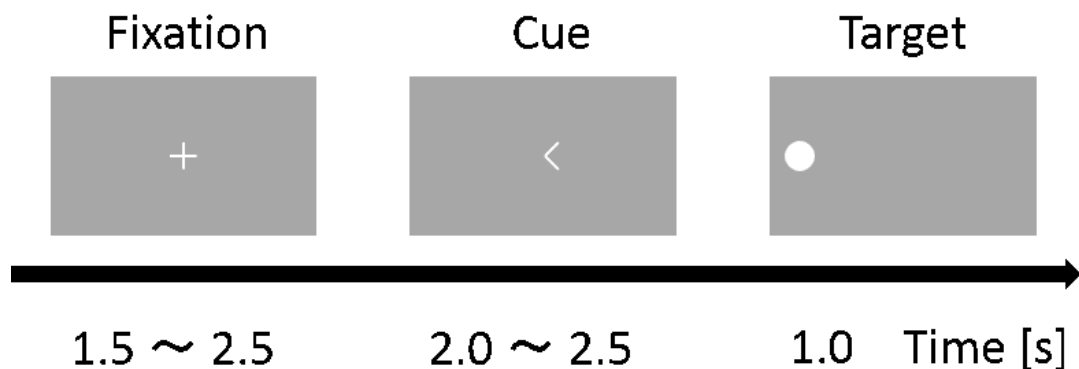


図 4.1: 課題の流れと画面表示例

円の大きさも変化させない。

MSC 評価区間である t_1 と t_2 については，予備実験を行い決定した．予備実験の結果，MSC の発生率の変化が顕著であったため， $t_1=0.7$ ， $t_2=0.9$ と設定した．MSC 検出手法には第 3 章で述べた手法を用いた．以上の課題を繰り返し，2 つの同心円の内側の円が外側の円と重なるまで大きくなったとき (円の拡大が 10 回発生したとき)，実験を終了した．

被験者には，Fixation と Cue が表示されている間は画面中央を注視し，表示された Cue をヒントに，Target が表示された瞬間，その表示方向とは逆方向にできるだけ早くサッカードするよう指示した．また，円が大きくなった課題が，良い注意状態と対応し，同心円の内側の円が外側の円と重なるまで大きくなると実験が終了することを伝え，できるだけ円が大きくなるように試行錯誤するよう指示した．ただし，被験者にはフィードバックの条件や MSC との関係は伝えなかった．

4.3.0.1 装置

本実験では，眼球運動計測装置として EyeSeeCam を使用した．Oculofeedback システムのプログラムは Mac Pro の OSX 上で動作させ，視覚刺激の表示には 29 インチのディスプレイ (LG 社製，型番 29UM67-P)，

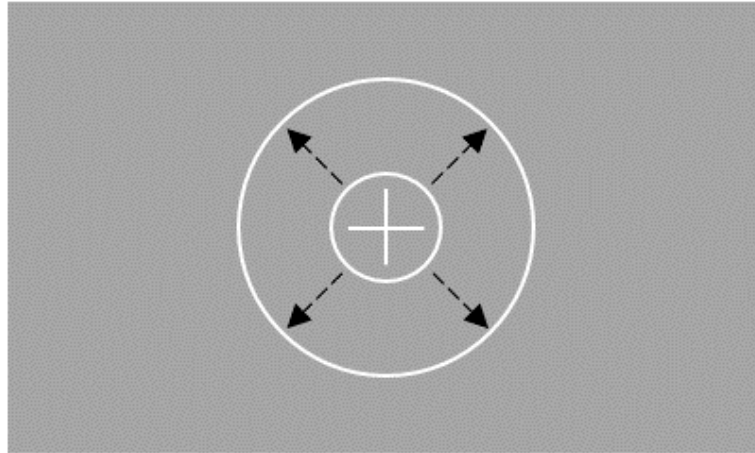


図 4.2: フィードバック画面

音の発生には Mac Pro 内蔵のスピーカーを用いた。

4.3.1 結果

6名の被験者が1日3回の実験に4日間（計12回の実験）参加した。6名の被験者のうち、3名はMSCによるフィードバックを与えるMSCFB群とし、他の3名は比較対象としてランダムなフィードバックを与えるRandFB群とした。

各被験者の実験結果を示す。以下、各図の左上には各被験者に対応する記号を付した。各図の上段の被験者はRandFB群3名の結果、下段の被験者はMSCFB群3名の結果である。

図 4.3 に実験回数に対する RT の推移を示す。各グラフの横軸は実験回数、縦軸は一回の実験の全課題の RT から推定した密度分布を基に算出した RT の最頻値である。なお、被験者 E に関しては1日目の実験回数を2回としたため、実験回数3における値はない。図 4.3 から RandFB 群の被験者 A に RT が遅れる傾向が確認できる。また、MSCFB 群の被験者 E に RT が早まる傾向が確認できる。

図 4.4 に実験回数に対する MSC 評価区間における MSC 発生率の推移を

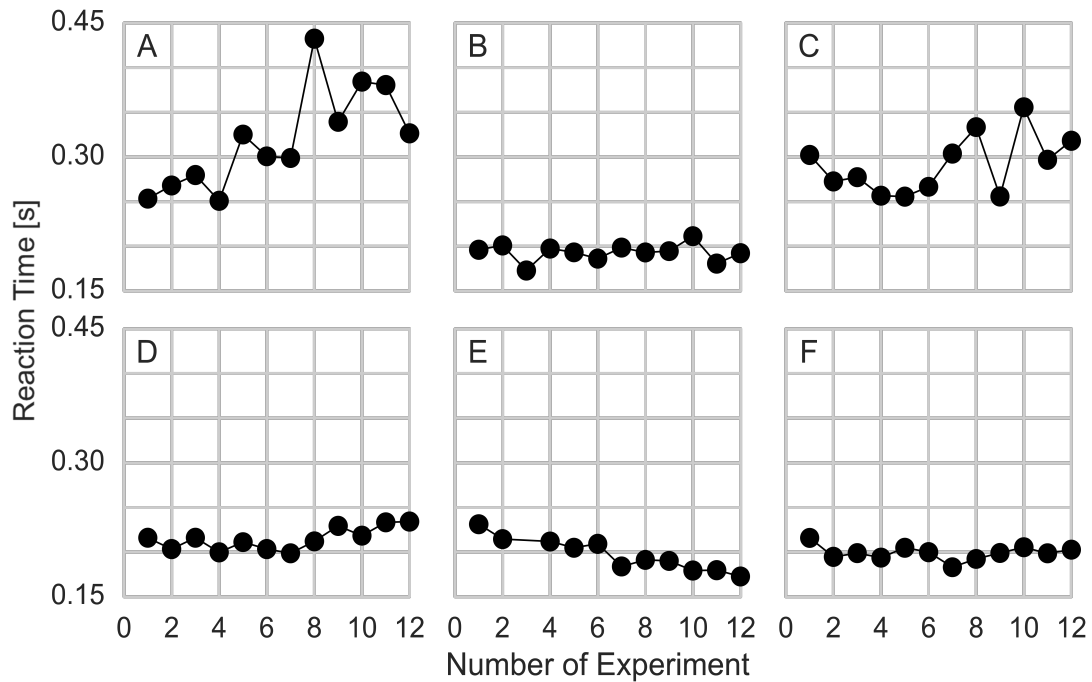


図 4.3: 実験回数に対するリアクションタイムの推移

示す．縦軸は 1 回の実験の全課題の MSC 評価区間における 1 秒当たりの MSC 発生率である．図 4.4 から RandFB 群の被験者 A に MSC 発生率の減少傾向が確認できる．また，MSCFB 群の被験者 D，F に MSC 発生率の増加傾向が確認できる．

図 4.5 に各被験者の課題中の MSC 発生率を示す．横軸は Cue の表示からの時間，黒い縦棒が Cue の表示された時間を表す．縦軸は 1 秒当たりの MSC 発生率である．図 4.5 から，全被験者において Cue の表示後 0.2 秒から 0.3 秒の間で MSC 発生率が顕著に増加していることが分かる．したがって，この区間を MSC Enhancement Interval (MSCEI) とし，区間内の MSC を解析した．

図 4.6 に実験回数に対する MSCEI における MSC 発生率の推移を示す．縦軸は一回の実験の全課題の MSCEI における一秒当たりの MSC 発生率である．図 4.6 から被験者 C，D，E，F で増加傾向が確認でき，特に被験者 E ではその傾向が顕著であることが分かる．

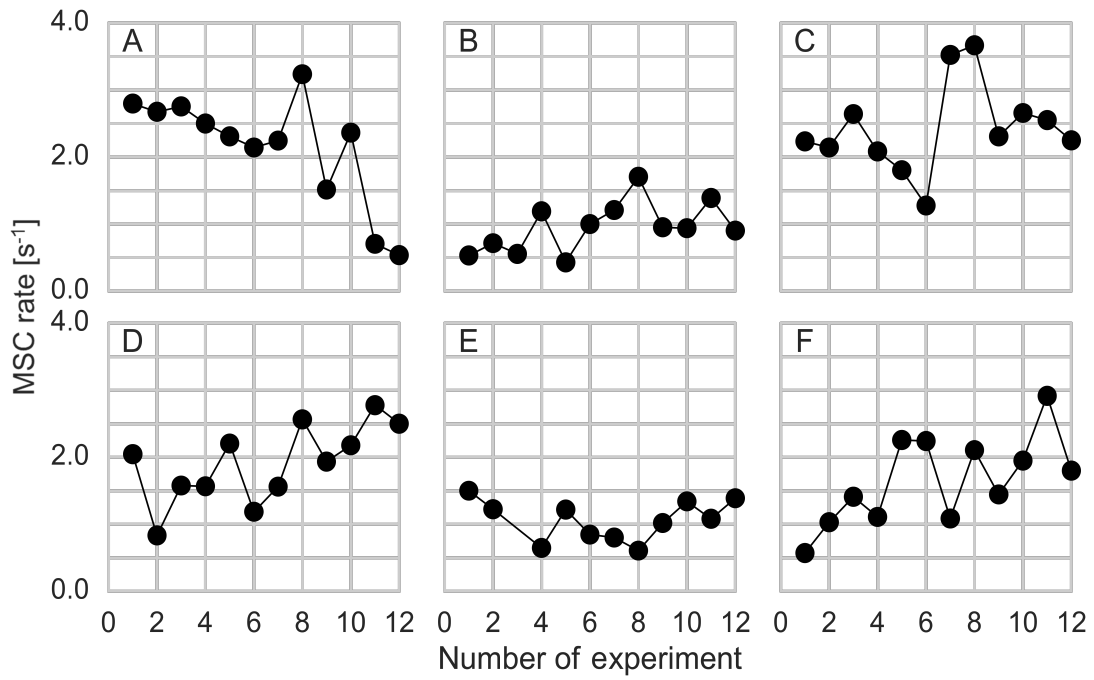


図 4.4: 実験回数に対する MSC 評価区間における MSC 発生率の推移

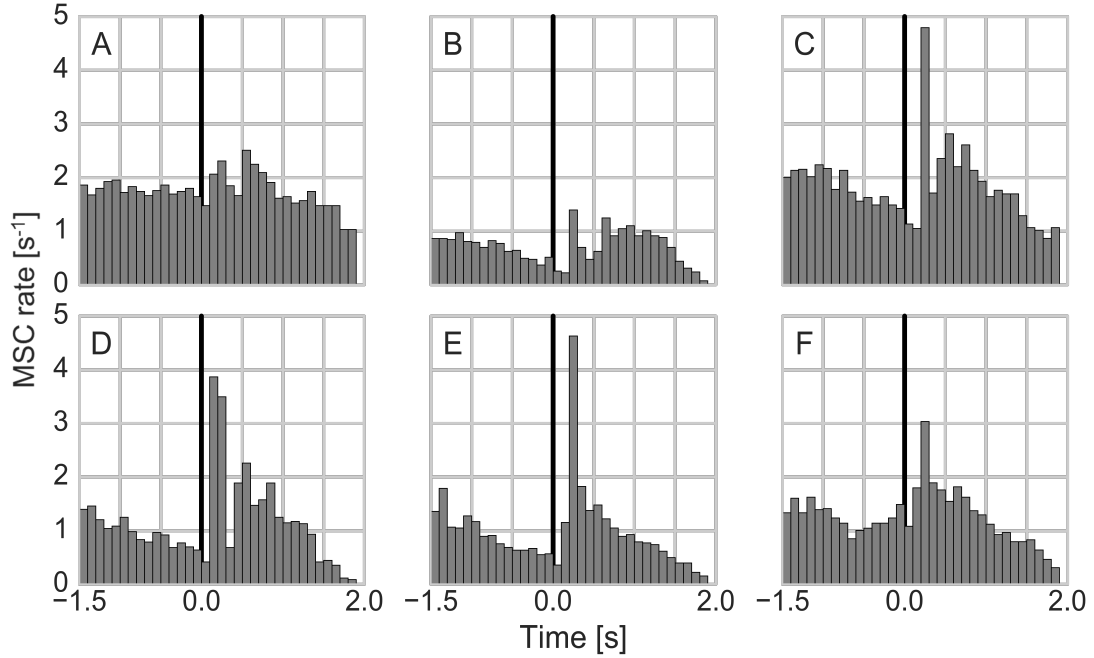


図 4.5: 各被験者のタスク中の MSC 発生率

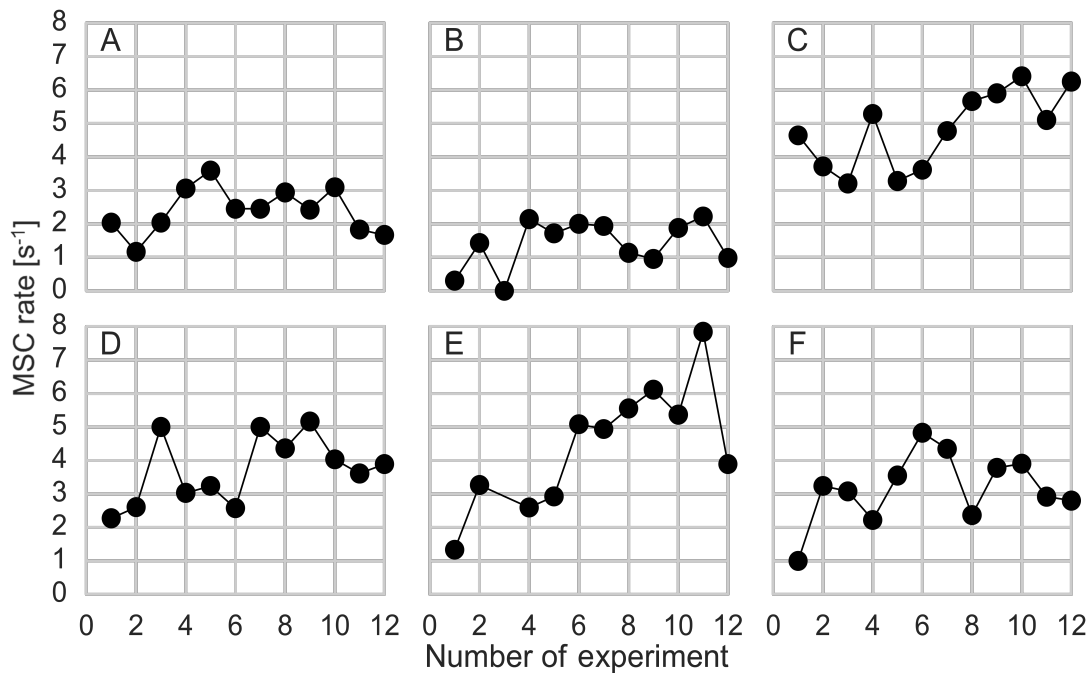


図 4.6: 実験回数に対する MSCEI における MSC 発生率の推移

Cue 表示直後である MSCEI で発生した MSC の発生方向は，Cue の方向との関連が考えられる．Cue と同方向に向けて発生した MSC を Cue congruent MSC (cc-MSCEI)，Cue と逆方向に向けて発生した MSC を Cue incongruent MSC (ci-MSCEI) として，MSCEI における MSC 方向の傾向を解析した．

図 4.7 に実験回数に対する MSCEI における MSC 中の cc-MSCEI および ci-MSCEI の割合の推移を示す．縦軸は 1 回の実験の全タスクの MSCEI で発生した MSC のうち cc-MSCEI であった割合（黒色）もしくは ci-MSCEI であった割合（灰色）である．図 4.7 から被験者 E 以外で MSCEI での cc-MSCEI の割合がほぼ 1 となることと，被験者 E で cc-MSCEI と ci-MSCEI の発生割合が等しくなる傾向が確認できる．

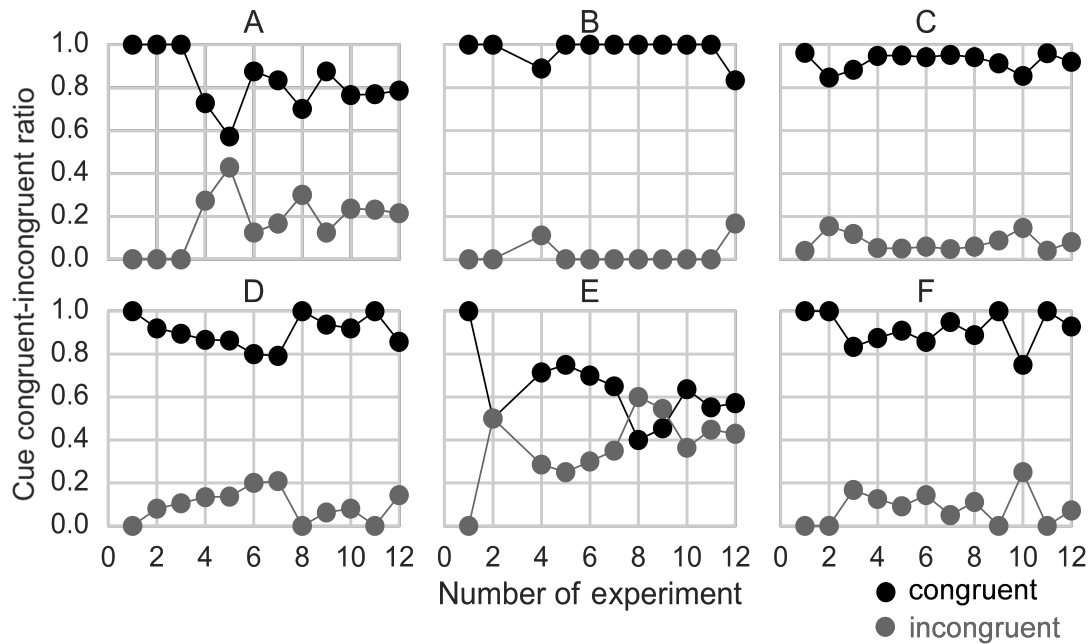


図 4.7: 実験回数に対する MSCEI における MSC 中の cc/ci-MSC の割合の推移

4.3.2 考察

開発した Oculofeedback システムを用い、MSC の発生傾向が課題中の注意状態に対応するものと仮定し、視聴覚刺激としてフィードバックすることで注意状態を操作する実験を実施した。フィードバックを実施した被験者のうち 3 名中 3 名（被験者 D, E, F）において、MSC 評価区間又は MSCEI 区間中の MSC が、実験回数に伴い増加する傾向が確認できた。また、その 3 名中 1 名（被験者 E）で、課題におけるリアクションタイムの改善と特徴的な MSC 発生傾向の変調が確認できた。この結果は、被験者 E が実験を重ねる中で課題に有効な注意の操作方法を学習し、それに伴う注意状態の変化が MSC の発生傾向に現れたものと考えられる。

一方、ランダムなフィードバックを与えた被験者群（被験者 A, B, C）で

は、一貫したリアクションタイムや MSC 傾向の変化は見られなかった。

以上の結果は、Oculofeedback には、MSC の発生特性の変調効果や実験中の潜在的注意の状態を改善する効果があることを示唆していると考えられる。

本実験では、先行研究の知見を元に実験パラダイムを設計することで、MSC の単純な発生傾向 (特定区間における MSC の有無) に基づき、期待される被験者の注意状態をフィードバックに利用した。このような実験において、より多くの被験者のパフォーマンス向上を誘発するには、最適な MSC 評価区間 (t_1, t_2) を探索するなど、各被験者の個人差に対応したフィードバック方法を模索し、実験パラダイムを工夫した上で、さらなる実験を実施する必要があると考えられる。

一方で、眼球運動には MSC 以外にサッカードなど、注意を反映するものがあり、本実験では、そのような情報を活用していない。Oculofeedback の効果を引き出すためには、眼球運動に含まれるリッチな情報を有効に活用し、機械学習などを用いて注意状態を推定することで、フィードバックを生成することが重要である可能性が考えられる。

第 5 章

畳み込みニューラルネットワークを応用した新たな眼球運動測定システム

第 4 章では，マイクロサッカードを利用して注意状態を操作する Oculofeedback 実験を実施したが，マイクロサッカード以外の様々な眼球運動にも，多様な脳の状態が反映されている．しかしながら，様々な眼球運動に含まれる多様な情報から，脳の状態推定に有効な特徴量を手作業で設計することは困難である．そこで，本章では，眼球画像から直接特徴量を抽出し，様々な脳の状態の推定に利用できる特徴抽出アルゴリズムを開発した．また，視線推定やサッカードなどの眼球運動イベント検出など，複数のタスクを同時に実行する手法とすることで，Oculofeedback に必要なデータを一括で得られるような眼球画像解析法を目指した．

5.1 Oculofeedback に適した眼球運動計測装置

本章では，小型演算装置を用いる Video Oculography (VOG) 方式の頭部装着型計測装置を想定した，新たな眼球画像の画像処理法を提案する．小型演算装置を用いた計測装置は，持ち運びが容易であり，屋外や動きのある実験における眼球運動計測や，Oculofeedback のように繰り返し実施する必要のある実験に適している．

VOG 方式の計測装置では，頭部に装着したカメラで撮影した眼球画像を画像処理することにより視線方向 (瞳孔中心) を推定する．近年，ディープラーニングを応用することで，高精度な視線推定が実現できることが示されている．例えば，DeepVOG[74] では，畳み込みニューラルネットワーク (ConvNet) を用い，約 0.5deg の精度でロバストに視線推定が可能である．

DeepVOG のようにディープラーニングを応用した従来法は，GPU を搭載した PC を使用することを前提としており，ConvNet のパラメータ数が多く，大きなメモリと計算量が必要となる．そのため，計算資源の限られる小型デバイスでは，リアルタイムに視線推定やイベント検出を実行することが困難である．また，DeepVOG では，視線方向の算出には ConvNet の出力の後処理が必要であり，まばたきや注視状態，視線移動 (サックード) などの眼球運動イベントの検出にはそれぞれに個別のアルゴリズムが必要である．

そこで，本章では，Oculofeedback に適した小型デバイスのための新しい画像処理法を提案する．提案法は，入力された眼球映像から，視線推定と眼球運動イベント検出の 2 つのタスクの出力を，end-to-end で一括処理するニューラルネットワーク (以下，モデルと呼ぶ) を用いる．2 つのタスクを同時に処理することにより，先行類似手法である DeepVOG で必要とされる独立したアルゴリズムによる後処理が不要となる．また，モデルの一部に MobileNetV2[75] を採用することで，ConvNet によるメモリの消費と計算

量を抑える。

本章では、まず、提案法において使用するモデルの構造とその学習方法を詳述する。次に、モデルの学習方法とその評価のためのデータセットの作成方法を説明し、作成したデータセットを用いた提案法の評価結果を示す。

5.2 提案法

5.2.1 概要

図 5.1 に提案法の全体図を示す。図左上の眼球映像 (input video) がモデルへの入力、図右下の特徴点 (output keypoints) と眼球運動イベント (output event) が出力である。白抜き矢印は時間の流れ、黒色矢印はデータの流れを表す。角丸の四角は黒色矢印による処理の内容を表す。

提案法のモデルは 3 つの畳み込みニューラルネットワーク (ConvNet) で構成される。これらを各々、特徴抽出部 (Feature Extraction)、特徴点検出部 (Keypoint Detection)、イベント検出部 (Event Detection) と呼ぶ。本モデルでは、まず、特徴抽出部が眼球映像の各フレームから特徴マップを算出し、フレーム数分の特徴マップの時系列 (sequence of feature maps) を出力する。特徴点検出部は特徴マップに基づいて特徴点を検出する。具体的には、眼球画像中の瞳孔中心点 (図右下 keypoint の画像上の赤点) と瞳孔楕円の短軸 (同緑点) ならびに長軸 (同青点) の端点各 2 点、及び目尻、目頭、上瞼と下瞼に沿う曲線の頂点の 4 点 (同橙点) の合計 9 点の特徴点の座標を出力する。イベント検出部は特徴マップの時系列から各フレームの眼球運動イベントを検出する。ここでは、眼球映像の各フレームが注視 (Fixation)、まばたき (Blink)、サッカード (Saccade) のうち、いずれの眼球運動イベント (図右下の event) に対応するものかを検出し、出力する。出力層の特徴次元数が検出する眼球運動イベントの数に対応し、提案法では出力層の特徴次元数は 3 である。検出する眼球運動イベントの種類は、対応するデータセット

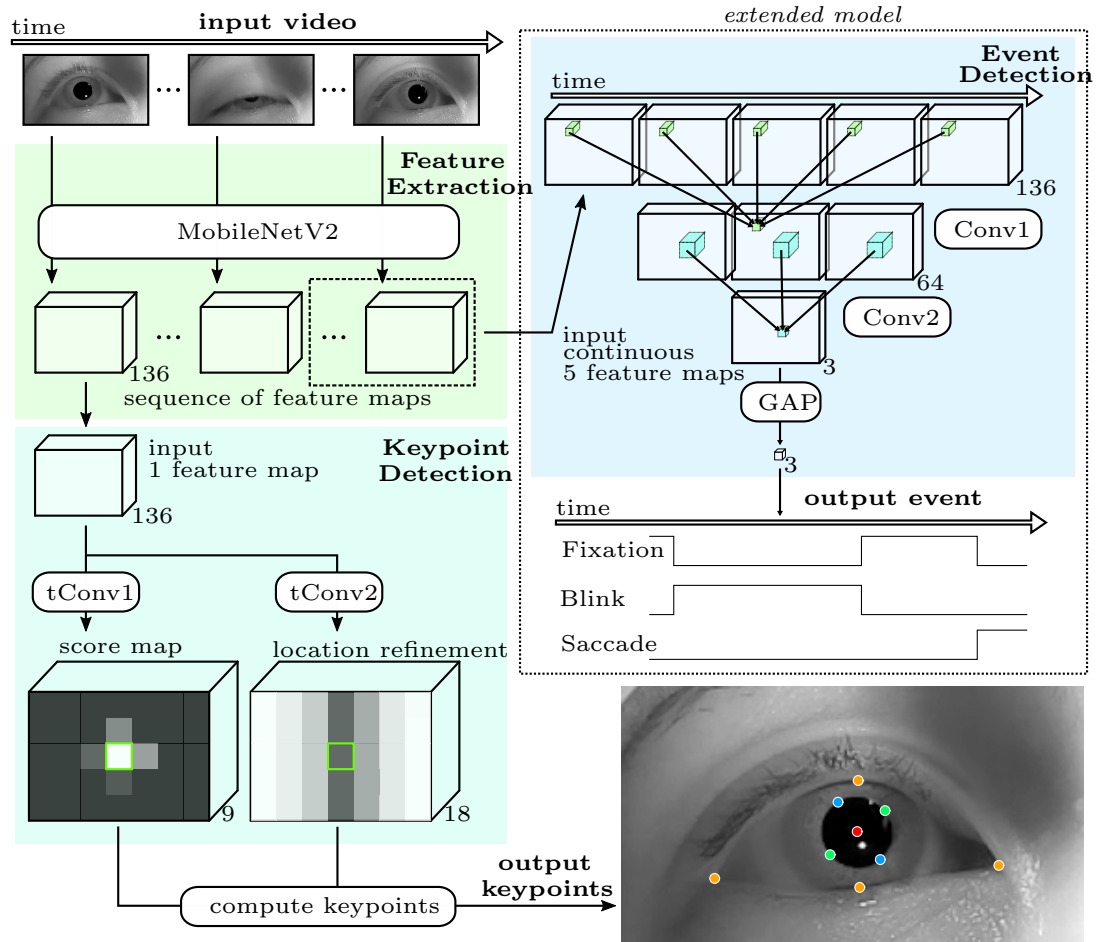


図 5.1: 提案法の全体図. 提案法のモデルは 3 つの ConvNet で構成され, それぞれ, 特徴抽出部 (Feature Extraction), 特徴点検出部 (Keypoint Detection), イベント検出部 (Event Detection) である. モデルは特徴点検出を学習した後, イベント検出部を追加し, イベント検出を転移学習する. 左上: 入力 of 眼球映像, 右下: 出力の特徴点と眼球運動イベント. 白抜き矢印は時間の流れ, 黒色矢印は入力から出力に至るまでのデータの流れ, 角丸の四角は黒色矢印による処理の内容を示す.

を追加し, 出力層の特徴次元数を増やすことで, 任意の数に拡張できる.

なお, 上述した検出結果については, 9 点の特徴点や 3 つの眼球運動イベント以外にも出力が考えられるが, 本研究では, 眼球運動の解析において広

く利用されている指標として，これらの特徴点と眼球運動イベントを検出することを選択した．また，提案法のモデルは，特徴抽出部の出力する特徴マップを利用して，2つのタスクを処理しているが，3つ以上の任意のタスクを処理するように拡張できる．例えば，眼球運動から眠気を推定する場合は，特徴マップを入力とした独立したニューラルネットワークを追加することで，本モデルの2つのタスクと同時に眠気推定をするモデルに拡張できる．

5.2.2 ConvNet の学習

ConvNet の学習には，一般に大量のデータが必要とされるが，提案法のモデルの学習・評価に利用できる公開された大量の眼球映像データセットは存在しないため，独自に学習・評価用のデータセットを作成する必要がある．しかしながら，大量の眼球映像を取得し，アノテーションを付与することは困難である．こうした状況において，DeepLabCut[76] は少量のアノテーション付き画像を用意するだけで，ヒトやマウス，馬など，様々な対象における高精度な特徴点検出を学習することに成功している．DeepLabCut は大量のデータで学習済みのモデルを転移学習させることにより，高精度な特徴点検出を実現していることから，本研究でも転移学習を用いることとした．転移学習は，学習済みのニューラルネットワークを，学習したものと異なるタスクやデータにおいて再利用するために，新たに用意したデータを用いて再学習する手法である．

DeepLabCut では，転移学習に ResNet[77] を利用しているが，提案法ではリアルタイムでの高精度な特徴点検出とイベント検出を実現するために，より高速な処理が可能で ResNet に匹敵する識別精度を誇る MobileNetV2[75] を利用する．MobileNetV2 に関しては，ImageNet[78] というデータセットで学習済みのものがインターネットで公開されている．

モデルの学習には，誤差逆伝搬法を用いたミニバッチ勾配降下法を用い

る．学習時にはモデルの入力画像の解像度に依存しない検出を実現するため，バッチごとに，バイリニア補間法によりランダムに拡大・縮小した画像を用いる．その際，特徴点座標も合わせて拡大・縮小する．

5.2.3 モデルの構造

上述の通り提案法のモデルは，特徴抽出部と特徴点検出部，イベント検出部の 3 つの ConvNet で構成される．以下，各 ConvNet の構造について詳細に述べる．

5.2.3.1 特徴抽出部

特徴抽出部の入力は眼球映像の 1 フレーム分の画像であり，出力は特徴マップである．学習済みの MobilenetV2 の出力を特徴抽出部の出力 (特徴マップ) とする．特徴マップの解像度は，入力画像の解像度を下げたものであり，その解像度は以下の様に算出する．まず，MobileNetV2 の畳み込み層には，ストライド s がある層とない層があり，ない層では入力画像の解像度は変わらないが，ある層では解像度を下げる．一般に，畳み込み層 1 層の入力画像の各辺の解像度を i とすると， i に対する出力解像度 o は，式 5.1 で計算できる [79]．

$$o = \lfloor \frac{i + 2p - k}{s} \rfloor + 1 \quad (5.1)$$

ここで，第 1 項のカッコは床関数で， p と k は，それぞれ一辺のパディング数とカーネルサイズである．

モデルには，14 層の MobileNetV2 を利用する．14 層目までの MobileNetV2 にはストライドがある層が 4 層あり，これらは $p = 1, k = 3, s = 2$ である．したがって，特徴マップの解像度は，式 5.1 を 4 層分繰り返し適用することで，入力解像度のおよそ $1/16$ となる．

また，各畳込み層ではモデルの表現力を高めるため次元数を 1.4 倍にす

る。したがって、出力の特徴次元数は、1次元の入力特徴次元数に対して136次元となる。

特徴抽出部の出力が、以下に述べる特徴点検出部とイベント検出部の入力となる。

5.2.3.2 特徴点検出部

特徴点検出部は DeepLabCut の出力層と同様に、1層の transposed convolution 層 (tConv 層)2 つで構成される。tConv 層によって出力の空間次元数は、入力された特徴マップの空間次元数の2倍となる。高精度な検出のためには連続値の特徴点座標が必要である。特徴点検出部の ConvNet では、score map と location refinement (locref) に基づいて特徴点の座標を算出する [80]。

図 5.1 の左下 (Keypoint Detection) に2つの tConv 層が特徴マップに基づいて出力する score map と locref 行列を示す。score map は、その空間次元の座標を入力画像の座標に対応付け、その座標に特徴点がある確率を表現する。特徴次元の1次元が特徴点1つに対応するため、検出する特徴点数が9である提案法では、特徴次元数は9次元である。特徴マップの空間次元数は、入力画像の空間次元数 (解像度) の $1/16$ 程度であることから、2倍しても元の空間次元数を復元せず、 $1/8$ 程度に縮小されている。したがって、score map は入力画像における特徴点のおおまかな座標を表していると言える。図 5.1 の score map は、枠が緑色の座標において、対応する特徴点 (ここでは瞳孔中心) が存在する確率が高いと推定した様子である。一方 locref は、特徴点の正確な座標を推定するために、score map と同様に対応付けた入力画像の座標と特徴点の座標までの距離を表現する。距離の縦と横の成分は2つの特徴次元でそれぞれ表す。1つの特徴点に2つの特徴次元が必要なため、提案法の locref の特徴次元数は18次元である。

図 5.1 左下の locref は、各座標の値が瞳孔中心からの横成分の距離を推定した様子である。緑色の枠が score map の最大値に対応する座標である。

各座標の明るさが特徴点の座標を補正する距離に対応する。緑枠に近いほど、補正に必要な距離が最も小さいことが分かる。特徴点の座標を算出するには、locref 行列の要素のうち score map が最大値を取った要素のみを使い、緑枠の座標にその locref の値を足し合わせる。

特徴点検出の損失関数として、DeepLabCut と同様に、score map には sigmoid 関数による交差エントロピー誤差を、locref には Huber 損失を用いる。

5.2.3.3 イベント検出部

イベント検出部は 2 層の畳み込み層と Global Average Pooling 層 (GAP 層) で構成される。図 5.1 の右上 (Event Detection) にイベント検出部の構造を示す。図中の箱の縦、横が空間次元で奥行きが特徴次元、白抜き矢印 (time) が時間次元を表す。イベント検出部の入力の特徴マップの時系列である。入力映像のフレームレートは 60FPS とする。イベント検出部は各フレームの前後 5 フレーム (約 83ms) の情報から、1 フレームごとの眼球運動イベントを推定する。そのために、まず、畳み込み層 1 層目 (conv1) では、空間次元 1 点で時間次元 5 点の畳み込み処理により、特徴次元数を 64 に削減する。次に、2 層目 (conv2) では、空間次元 3 点で時間次元 3 点の畳み込み処理により、特徴次元数をイベント数まで削減する。そして、GAP 層では、空間次元において特徴次元ごとに平均をとり、各イベントの確率を算出する。イベント検出の損失関数として、softmax 関数による交差エントロピー誤差を用いる。

図 5.1 右下の output event は、上述の処理を入力映像の全フレームに対して実行し、出力値が最大となったイベントを 1、その他を 0 として時系列表示したイベント検出結果の例である。

提案法のモデルは、畳み込み層や GAP 層など空間次元に依存しない処理で構成されるため、入力画像の解像度には依存しない構造であり、任意の解像度の入力画像を用いることができる。また、特徴抽出部と特徴点検出部に

については、時間に依存しない処理で構成されるため、特徴点の検出に限っては、入力眼球画像 1 枚から処理できる。

5.3 評価方法

本評価では、提案法の有効性を吟味するために、次の 5 項目に着目する。すなわち、(1) 転移学習した特徴抽出部が抽出した特徴マップには特徴点検出と眼球運動イベント検出に有効な特徴が反映されていること、(2) 特徴マップを利用することで特徴点検出が実現できること、そして同時に、(3) 特徴点検出で学習済みの特徴マップを利用してイベント検出も実現できること、また、(4) 提案法が異なる装置や被験者のデータにも適用でき、(5) 眼球運動に関連するマルチタスク学習に拡張可能であること。以下、評価のためのデータセット取得のために用いた眼球運動計測装置と評価手順について記す。

5.3.1 眼球運動計測装置

本評価では、既存の VOG 方式の眼球運動計測装置によって推定された眼球特徴点の座標を提案法の教師データとして利用する。ここでは、計測装置として EyeSeeCam (EyeSeeTech 社製)[16] を使用した。EyeSeeCam は医療用にも広く用いられており、最大 500fps で両眼の眼球映像と共に、高精度に眼球回転角度 (yaw ならびに pitch) を記録できる。また、異なる眼球運動計測装置により撮影された眼球映像へ適用した場合の性能も評価するため、Raspberry Pi Zero W に赤外線 LED 付きのカメラ (kuman 社製、型番: SC15) を組み合わせて頭部に固定する眼球運動撮影装置 (以下、RPiCam) を自作した。RPiCam は撮影した映像を WiFi で無線送信する機能のみを有し、画像処理は映像を受信したノート PC などで行う。

5.3.2 手順

上記 (1)～(5) の項目を評価するために、本章では以下の手順で提案法の ConvNet を 4 回転移学習する。

まず、提案法の特徴点検出とイベント検出は、関連するタスクではあるものの時間スケールが異なるため、提案法のモデルを特徴点検出とイベント検出で個別に学習して評価する。すなわち、特徴点検出では、視線方向の異なる眼球画像を 1 フレームごとに学習すればよいが、イベント検出では、発生している眼球運動の異なる映像ごとに学習する必要があるためである。したがって、学習のタイムスケールの異なる 2 つのタスクを同時に実行することによる実装の複雑さを避けるため、学習手順を 2 つのタスクが独立に転移学習するようにした。

図 5.2 に転移学習の手順を示す。図中の長方形は矢印の前までの ConvNet を転移学習することを示す。長方形の中には、学習内容と利用する被験者群 (A, B, C) とデータセットの名前 (dev1, dev2, dev3, dev4) を示した。白色の長方形は特徴点検出の転移学習であり、灰色の長方形はイベント検出の転移学習である。図の最下段の角丸の図形は試験用のデータセット (test1, test2) による 2 つのタスクの試験である。本評価では、3 つのデータセットで特徴点検出を学習し、1 つのデータセットでイベント検出を学習した後、各タスクを試験用のデータセットで試験する。

各データセットは 3 つに分けた被験者群 (A, B, C) のデータから作成する。被験者群 A のデータは、EyeSeeCam と RPiCam の両方の装置で 1 回ずつ計測し、両装置における提案法の特徴点検出の評価に使用する (具体的なデータ数は 5.4.1 項で述べる)。また、RPiCam を使用した場合のイベント検出の評価にも使用する。被験者群 B のデータは RPiCam で計測し、被験者群 A のデータで学習した ConvNet を転移学習することにより、異なる被験者から取得したデータにおける提案法の評価に使用する。被験者群 C

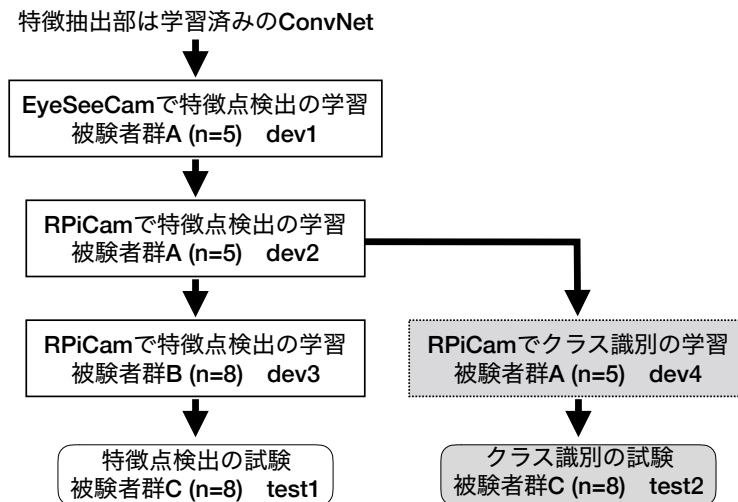


図 5.2: 学習と評価の手順

のデータは ConvNet の学習には用いず，特徴点検出及びイベント検出を学習した ConvNet の評価にのみ使用する．

5.4 データセット

5.3.2 項で述べた 4 回の学習と 2 回の試験のために，6 種類のデータセットを作成する．そのためにまず，眼球映像を記録する実験を実施した．次に，記録した眼球映像から特徴点検出とイベント検出用のデータをそれぞれ抽出した．そして，抽出した各データにアノテーションを付与した．

5.4.1 実験

データセット作成のために 21 名の被験者から眼球運動を記録した．以下，その手順を述べる．

被験者は暗室で椅子に座り，ディスプレイ (35-inch, BENQ 社製 MARS3500-B) から 62cm 離れた距離で顎台にあごを載せて頭部を安定さ

せた．ディスプレイのサイズは横 67deg, 縦 31deg である．被験者の注視点として, 中央に白い点 1 点と, それを中心とした半径 3, 5, 8.5, 12, 16, 20, 24, 28, 32, 36deg の 10 個の同心円に, それぞれ 8, 12, 16, 20, 24, 28, 32, 36, 40, 44 点を等間隔に配置した白い点を生成した．生成した点の内, ディスプレイの範囲内にある合計 135 点を表示した．

被験者は指定された順に各点を 1 回ずつ注視した．注視すべき点は, 円で囲うことにより指定した．この円はランダムに動かされ, 135 点を一回ずつ約 1 秒間囲うようにした．このときの眼球映像を約 2 分 30 秒間に亘り 1 本の映像として記録した．この眼球映像には, 被験者の眼球運動イベントとして, ”視線移動 (サッカード)”, ”注視”, ならびに自然な”まばたき”が含まれている．

21 名のうち 5 名を被験者群 A, 残りの 16 名を 8 名ずつ被験者群 B, C とした．被験者群 A の 5 名には EyeSeeCam を装着してもらい 256fps で, すべての被験者 21 名には RPiCam を装着してもらい 60fps で, それぞれ 1 回ずつ, 合計 26 本の眼球運動中の映像を記録した．

EyeSeeCam で記録した映像の解像度は幅 188, 高さ 100 ピクセルである．一方, RPiCam で記録した映像は, 被験者ごとに手動で領域を設定し, 図 5.1 の右下の画像のように切り出した．したがって, 解像度は映像ごとに異なり, 本実験では, 幅は 180 から 270, 高さは 85 から 150 の範囲である．

5.4.2 データの抽出

EyeSeeCam と RPiCam で記録した眼球映像から画像又は映像を切り出し, 特徴点検出とイベント検出の学習及び評価用のデータセットをそれぞれ作成した．表 5.1 に抽出した画像枚数と映像数, データセットの名前をまとめた．

表 5.1: データセットの一覧

装置	取得した映像		特徴点検出		イベント検出の映像数 ¹	
	被験者群	被験者数 (映像数)	名前	画像数	名前	まばたき サッカード
EyeSeeCam	A	5	dev1	500		
	A	5	dev2	250	dev4	50 100
	B	8	dev3	400		
	C	8	test1	400	test2	80 160

¹ 注視はまばたき及びサッカード中の映像に含まれる.

5.4.2.1 特徴点検出

このタスクでは，様々な方向へ視線を向けたときの画像が必要である．提案法のモデルは，特徴点検出においては時間に依存しないため，視線方向が同様の画像が多く含まれてしまう映像ではなく，画像により特徴点検出用のデータセットを構成した．

データセット用の画像として，EyeSeeCam で記録した被験者群 A では，5 本の映像から注視中の画像を各 100 枚ずつ，実験開始後から 1 秒間隔で抽出して合計 500 枚の画像を，RPiCam で記録した被験者群 A，B，C では，それぞれ 5，8，8 本の映像から同様に各 50 枚ずつ，それぞれ合計 250，400，400 枚の画像を抽出した．

5.4.2.2 イベント検出

イベント検出部は，時間に依存した処理が必要なため，イベント検出用のデータセットは映像で構成した．

データセット用の映像として，RPiCam で記録した被験者群 A，C の映像から，まばたきとサッカードが発生している区間を目視で確認し，まばたきが発生している映像 10 本，サッカードが発生している映像 20 本を，被験者毎にランダムに抽出した．したがって，被験者群 A の 5 名から，まばたき映像 50 本，サッカード映像 100 本，被験者群 C の 8 名から，まばたき映像 80 本，サッカード映像 160 本を抽出した．なお，注視はまばたき及びサッカード中の映像に含まれる．抽出した映像の長さは，注視，まばたき，サッカードの検出に十分な長さとして，16 フレーム (約 0.27 秒) とした．

5.4.3 アノテーション

抽出したデータには以下のようにアノテーションを付与した．

5.4.3.1 特徴点検出

特徴点検出用に抽出した画像に，目視で 9 つの特徴点の座標をアノテーションとして付与した．EyeSeeCam で記録した画像の瞳孔中心点のアノテーションについては，EyeSeeCam が出力したものを活用した．アノテーションには，独自に開発した GUI アプリケーションを用いた．開発したアプリケーションは，眼球画像を見ながら，マウスを用いて眼球画像上に特徴点の座標を付与できる．瞳孔中心とその周辺の特徴点 5 点の座標については，楕円の中心点と長軸，短軸の長さ，回転角度の 3 つのパラメータを調整することにより楕円を瞳孔の縁に当てはめるようにアノテーションを付与した．残りの眼の形状の特徴点 4 点の座標については，目視で特徴点を確認し，眼球画像上にマウスで座標を指定することでアノテーションを付与した．

5.4.3.2 イベント検出

イベント検出用に抽出した映像の 1 フレームごとに，映像を再生しながら目視で，発生中の眼球運動イベント（注視，まばたき，サッカード）をアノテーションとして付与した．映像中で，まばたきもサッカードも発生していないフレームのアノテーションは，被験者に課したタスクの性質から全て注視とした．

アノテーションの内訳に関しては，まばたき映像では，16 フレームほぼすべてが，まばたき中のフレームであった．サッカード映像では，16 フレーム中 4～12 フレーム程度が，サッカード中のフレームであった．両映像のまばたき，サッカード発生中以外のフレームが注視中のフレームである．

5.5 評価結果

提案法の ConvNet を図 5.2 の手順に従って学習させた．5.2.2 項の学習方法で述べたように，学習時には，入力画像を 0.64 倍から 1.5 倍にランダムに拡大・縮小し，学習率は 0.005 とした．特徴点検出とイベント検出の評価結果を以下に述べる．

5.5.1 特徴点検出

5.5.1.1 学習

学習済み MobileNetV2 を特徴抽出部として，データセット dev1 で転移学習した．特徴点検出の学習では，特徴抽出部と特徴点検出部の ConvNet のパラメータを更新する．データセット dev1 の 500 枚の画像を被験者ごとに 8:2 に分け，それぞれ学習データ (400 枚) と検証データ (100 枚) とした．図 5.3 の最上段に学習曲線を示す．横軸のエポック数に対して縦軸が損失である．黒い曲線が学習データ，灰色の曲線が dev1 の被験者 5 名それぞれの検証データにおける学習曲線である．学習曲線から 1000 エポックで学習が収束したと判断し，1000 エポック学習直後の ConvNet を dev1 による学習済み ConvNet とした．

データセット dev1 による，学習済みの ConvNet を，dev2 で転移学習した．データセット dev1 と同様に，dev2 の 250 枚のうち 200 枚を学習データ，残りの 50 枚を検証データとした．図 5.3 の中段が dev2 の学習曲線である．dev1 での学習と同様に学習曲線から 1000 エポックで学習が収束したと判断し，1000 エポック学習直後の ConvNet を dev2 による学習済み ConvNet とした．

データセット dev2 による，学習済みの ConvNet を，dev3 で転移学習した．上述した 2 回の学習と同様に，dev3 の 400 枚のうち 320 枚を学習デー

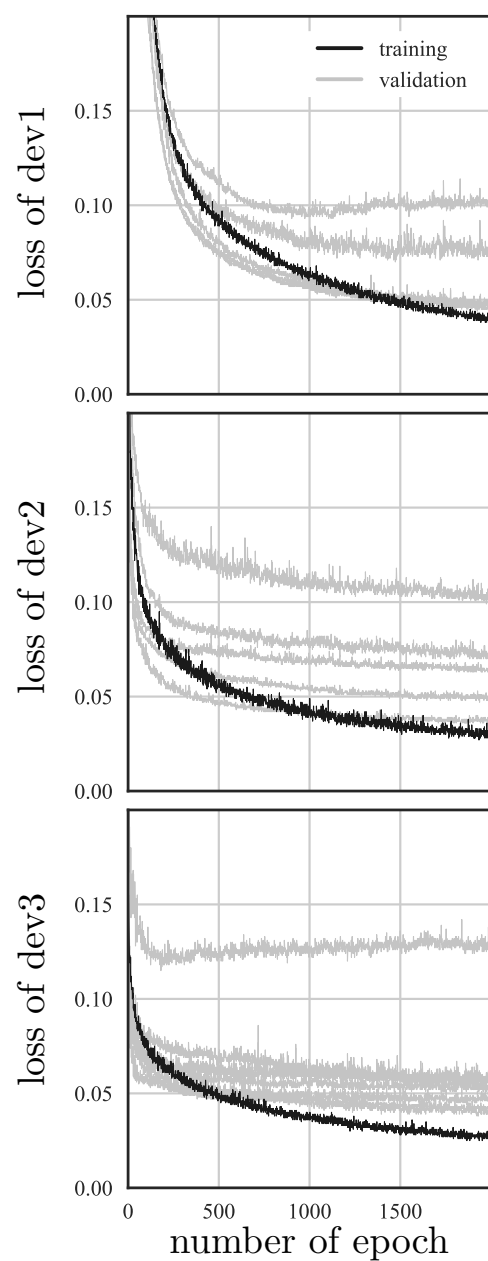


図 5.3: 転移学習ごとの学習曲線. 上段は dev1, 中段は dev2, 下段は dev3 使用時の学習曲線.

タ, 残りの 80 枚を検証データとした. 図 5.3 の最下段が dev3 の学習曲線である. 学習曲線から 200 エポックで学習が収束したと判断し, 200 エポッ

ク学習直後の ConvNet を dev3 による学習済み ConvNet とした。

5.5.1.2 精度

アノテーションの特徴点の座標と提案法により推定された対応する特徴点の座標の Root Mean Square Error (RMSE) を推定誤差として、提案法による特徴点座標の推定精度を評価した。

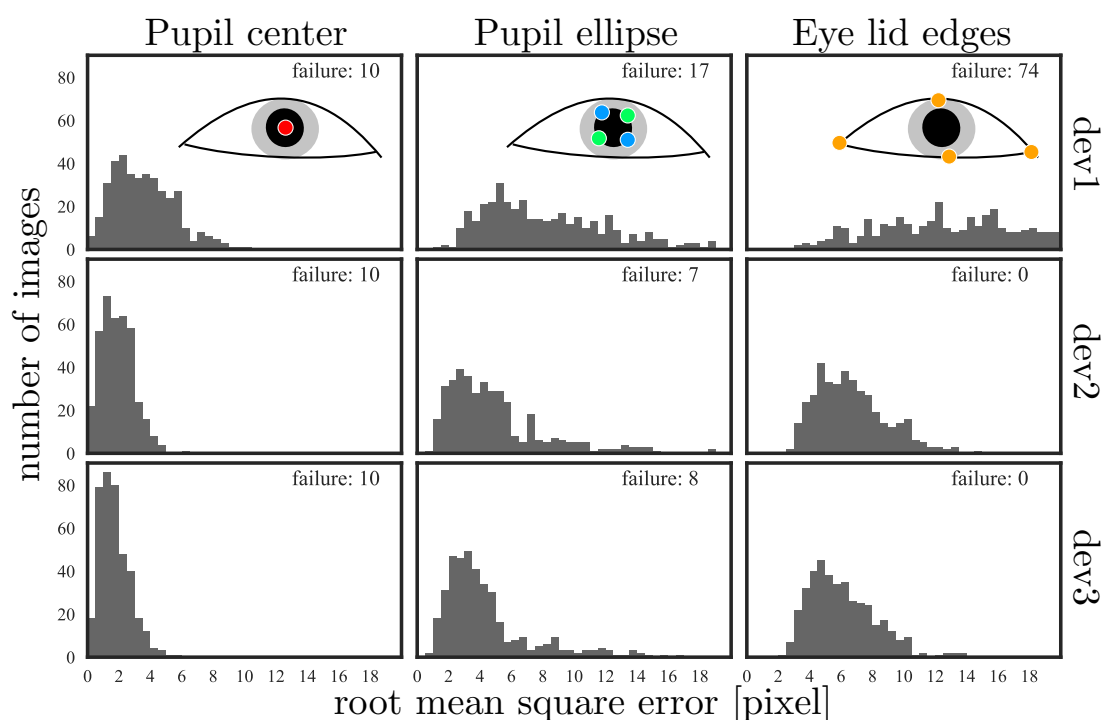


図 5.4: 転移学習ごとの瞳孔中心 (左) と瞳孔楕円 (中央), 眼の形状 (右) の特徴点検出の誤差のヒストグラム. 上段は dev1, 中段は dev2, 下段は dev3 使用時のヒストグラム.

図 5.4 は試験用のデータセット test1 において推定した画像ごとの特徴点座標の推定誤差のヒストグラムで、横軸が RMSE, 縦軸が画像枚数である. 各行は使用したデータセットであり, 最上段から順に dev1, dev2, dev3 により学習済みの ConvNet で推定した結果である. 各列は 9 点の特徴点を種類ごとに 3 つに分けたものであり, 左から順に, 瞳孔中心 (1 点), 瞳孔楕円

表 5.2: 特徴点検出の精度

データセット	推定誤差 (平均 $\pm$ 標準偏差 ピクセル)		
	瞳孔中心	瞳孔楕円	眼の形状
dev1	4.09 $\pm$ 5.04	9.52 $\pm$ 10.5	16.9 $\pm$ 12.8
dev2	2.35 $\pm$ 3.87	5.48 $\pm$ 5.88	6.68 $\pm$ 4.29
dev3	2.14 $\pm$ 4.19	5.03 $\pm$ 6.19	6.08 $\pm$ 4.19

(短軸と長軸の端点各 2 点計 4 点), 眼の形状 (目尻, 目頭, 上瞼, 下瞼の 4 点) である. ヒストグラムの横軸は 0 から 20 ピクセルまでとして, それ以外の推定結果は, 推定失敗とした. 各ヒストグラムの右上に示した値は 400 枚中推定に失敗した画像の枚数である. 推定に失敗した画像は, 主にまばたき中のものであった.

表 5.2 の各行に, 瞳孔中心, 瞳孔楕円, 眼の形状の, 各データセットにおける学習後の推定誤差の平均値と標準偏差を示す. なお, 眼球画像の大きさは平均でおよそ縦横 100x200 ピクセルである.

5.5.2 イベント検出

5.5.2.1 学習

被験者群 A の特徴点検出用のデータセット dev2 による, 学習済みの ConvNet を, イベント検出用のデータセット (dev4) を使用して, イベント検出部を構成する ConvNet を転移学習させた. イベント検出の学習では, 特徴抽出部のパラメータは固定とし, イベント検出部のパラメータのみを更新した. dev4 の 150 本の映像のうち, 120 本を学習データ, 30 本を検証データとした.

図 5.5 に学習曲線を示す. 横軸のエポック数に対して縦軸が損失である.

黒い曲線が学習データ，灰色の曲線が検証データにおける学習曲線である．学習曲線から学習エポック数 1000 回を学習の収束と判断し，イベント検出精度を試験用データセット (test2) で評価した．

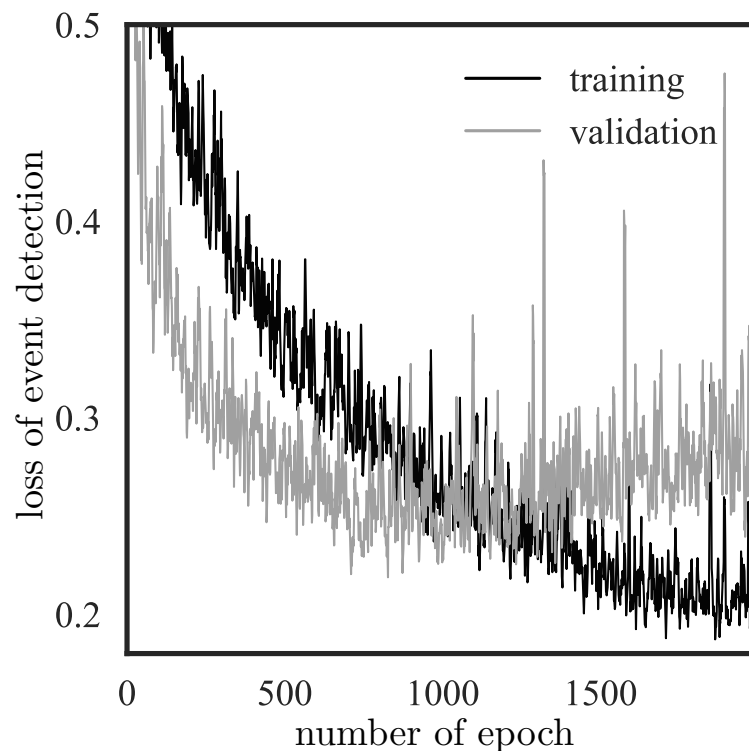


図 5.5: イベント検出の学習曲線

5.5.2.2 精度

表 5.3 に各眼球運動イベントの検出精度と F1 スコアを示す．検出精度は，test2 のまばたき映像 80 本とサッカード映像 160 本，計 240 本の 16 フレームの映像に含まれる，まばたき，注視，サッカード中の 3840 フレームの内，眼球運動イベントを正しく推定できた割合とした．サッカードの識別精度が他の眼球運動イベントより低い，これはサッカード発生中のフレーム数が他の眼球運動イベントのフレーム数に対して相対的に少なく，十分に学習が進まなかったためであると考えられる．test2 の各映像中でサッカードが発

表 5.3: イベント検出の精度

	検出数 / 合計イベント数	識別精度	F1 score
まばたき	1008 / 1165	86.5%	0.839
注視	1455 / 1747	83.3%	0.807
サッカード	568 / 928	61.2%	0.681

生じたかどうかだけを評価すると，検出率は 87.0%(F1 スコアは 0.906) となった．

5.6 考察

第 5.3 節で述べた 5 つの評価項目について，評価結果に基づいて考察する．

5.6.1 特徴点検出

図 5.4 の最上段は，EyeSeeCam を用いたデータセット dev1 で学習した ConvNet の推定誤差を，RPiCam を用いたデータセット test1 において算出したものである．このときの瞳孔中心 (左) の推定誤差は，瞳孔楕円 (中央) と眼の形状 (右) の推定誤差に比べて小さい．また，EyeSeeCam と RPiCam で用いた各々性能の異なるカメラは，画角や解像度が異なる眼球画像を撮影した．すなわち，提案法のモデルは，学習に使用しなかった RPiCam で記録された眼球画像の特徴点検出が可能であり，眼球撮影装置を変えた場合でも特徴点検出が可能であると期待できる．

また，図 5.4 と表 5.2 において，中段の dev2 と最下段の dev3 では被験者を追加して学習しても推定精度にほとんど変化がないことから，提案法は異なる被験者に対しても高い汎化性能を有していると考えられる．

提案法は瞳孔中心座標をおよそ 2pixel の誤差で推定できることから、視線推定についてはマーケティングやゲーム用途などで充分に応用可能な精度を実現したと考えられる。

一方、図 5.3 の検証データにおける学習曲線を見ると、被験者により収束する損失値が異なっていることから、単純に被験者を増やすのではなく、映像ごとの画像の質や個人差を考慮してデータを追加する必要がある事がわかる。

また、図 5.4 の各段のヒストグラムを見ると、dev2 から dev3 への転移学習の推定誤差の改善は、dev1 から dev2 への転移学習に比べて小さいことから、より高精度化を図るためには、データの追加だけでなく、特徴点間の依存関係を考慮したモデルへの改良など、ConvNet の構造についても改良が必要と考えられる。

第 5.3 節で述べた評価項目については、(2) 視線推定のための特徴点検出が実現でき、(4) 装置や被験者が異なる場合でも汎化性能を有することから、異なる装置や被験者に適用できることが示されたと考えられる。

5.6.2 イベント検出

表 5.3 を見ると、F1 スコアで 0.8 以上の値が得られたことから、提案法により、眼球運動のイベント検出が高精度で可能であると言える。

また、イベント検出の学習では特徴抽出部のパラメータを更新しなかったことから、イベント検出の結果は特徴点検出と同一の特徴量に基づいていることになる。このことから、特徴点検出で学習した特徴抽出部の出力した特徴マップには、特徴点検出だけでなく、眼球運動イベントの推定にも有効な特徴が反映されていると考えられる。

評価項目については、(1) 学習した ConvNet の特徴マップには特徴点検出とイベント検出に有効な特徴が反映され、(3) その特徴マップを利用してイベント検出が可能であることを示したと考えられる。また、イベント検出

部は、視線推定タスクを学習後に追加した部分であり、眼球運動イベントの検出タスクを拡張したものであることから、(5) 眼球運動に関連するマルチタスク学習へ拡張できることを示せたと考えられる。

5.6.3 精度と軽量化のトレードオフ

DeepVOG と提案法で採用しているニューラルネットワークの規模を、パラメータ数を指標として比較した。まず、DeepVOG のパラメータ数が約 3700 万であるのに対して、提案手法のパラメータ数は約 170 万であった。一方、DeepVOG の視点推定の誤差が、 320×240 の画像解像度において 1 ピクセル程度であるのに対して、提案法の誤差は平均 200×100 の画像解像度において 2 ピクセル程度である。このようなトレードオフから、両手法は用途に応じて使い分ける必要があると考えられる。

5.6.4 学習手順

本研究では、5つの評価項目を評価するために多段階転移学習を実施した。多段階転移学習では、不適切な転移元による性能の悪化が報告されている [81]。図 5.2 の結果については、特に dev1 の Eye lid edges が適合していないため、Eye lid edges については、より適切な転移元を選択する必要があると考えられる。

また、学習手順は、適用先のデバイスによって適切に設計する必要があると考えられる。特に、用意したデバイスで大量の学習用データが確保できる場合には、転移学習を利用せず、ランダムなパラメータから学習を始めることも検討すべきである。一方、用意したデバイスで十分なデータを得ることが困難な場合には、EyeSeeCam のような既存の眼球運動計測装置を利用して学習データを用意し、多段階転移学習することで、効率的に高精度な推定モデルが得られると考えられる。

図 5.2 の結果から，Pupil center については，多段階転移学習により自然画像で学習済みのモデルを，高精度に視線推定が可能な眼球運動計測装置を利用してあらかじめ転移学習しておくことで，新しいデバイスでは少量のデータで転移学習できることが示されたことから，提案法は簡易に新たな眼球運動計測装置を実現できる手法であると言える．

5.7 まとめ

本章では，ディープラーニングを応用し，VOG 方式の眼球運動計測装置のための新しい画像処理手法を提案した．画像処理に用いる ConvNet に MobileNetV2 を採用し，視線推定とイベント検出を End-to-end で一括処理することで，メモリの消費量と計算量を抑えたモデルを開発した．提案法の有効性を示すために，独自に評価方法を設計し，既存の高性能な眼球運動計測装置に加え，安価な眼球運動計測装置を自作して，21 名の被験者からデータセットを作成した．そして，作成したデータセットを用いてモデルを学習させて評価した．その結果，提案法は装置や被験者に依存しない汎化性能を有しており，マルチタスク処理に拡張可能であることを示した．

提案法は，眼球画像から眼球運動の解析に有効な特徴量を抽出する ConvNet を獲得した．提案法のモデルは，目的に合わせて 3 つ以上のタスクの処理に容易に拡張できるものであり，多様な用途への応用が期待できる．例えば，本研究では扱わなかった滑動性眼球運動 (smooth pursuit)，視運動性眼球運動 (optokinetic response)，前庭動眼反射 (vestibulo-ocular reflex)，マイクロサッカード (microsaccade) の検出などのために，眼球映像を end-to-end で処理する眼球運動解析システムの構築に役立つと考えられる．

近年ディープラーニング関連技術の開発が急速に進んでおり，GPU だけでなく，スマートフォンや組み込みデバイス向けにも，TPU などニューラルネットワークによる推論を高速に実行可能なプロセッサが開発され，市場に

投入されている．提案法のように，眼球運動解析を ConvNet で end-to-end に一括して処理する手法は，従来法と異なり，これらの技術の恩恵を受けることができる．

第 6 章

総括

6.1 本論文の成果

神経科学や機械学習の進展により，これまで実現困難であった眼球運動の計測や人の状態推定などの高度な情報処理が可能になってきた．本研究では，このような状況を受け，Oculofeedback を提案し眼球運動データの新たな活用法を示した．

本論文の成果は以下のとおりである．

Oculofeedback の提案とそれを実現するシステムの開発

被験者に眼球運動に基づく脳の状態をフィードバックする Oculofeedback を提案した．また，その実現のために必要な Oculofeedback システムを実装した．

これまでにない実験系を提供することは，脳の機能に関する新たな知見の発見や脳の機能を改善・向上させる画期的な訓練法の開発に繋がる貢献であると考えられる．

実時間高精度マイクロサッカード検出法の開発

眼球運動時系列波形からマイクロサッカードをロバスト・高精度に検出する手法を開発した。

マイクロサッカードは、眼球運動の中でも役割・機能に未解明の部分があり、研究が活発な分野である。従来よりもロバストで高精度なマイクロサッカード検出を実時間で実行できる手法には需要があるため、マイクロサッカード研究の発展に寄与すると考えられる。

マイクロサッカードを利用した注意状態を操作する Oculofeedback 実験

被験者の注意状態を操作することで特定の課題におけるパフォーマンスの向上を確認した。実験において、課題のパフォーマンスであるリアクションタイムの改善とそれに伴う眼球運動の変化が確認できた。実験結果から、Oculofeedback には、被験者が注意を操作する能力を訓練する効果があることが示唆された。

ヒトの注意状態を操作し、課題に最適な注意状態を引き出すことで注意力を向上させる訓練は、工場や工事現場で働く作業員やアスリート、自動車の運転者など、注意力の維持やコントロールが重要となる人々のパフォーマンス向上に繋がる将来性の期待できる画期的な手法であると考えられる。

新しい眼球運動測定システムの開発

Oculofeedback に最適な眼球運動測定システムの実現を目指し、畳み込みニューラルネットワークを用いた新しい眼球画像処理法を開発した。開発したアルゴリズムは、視線推定と同時にまばたきやサッカードなどの眼球運動イベントも同時に検出可能な手法である。今後は、入力画像のノイズレベル

や Oculofeedback のためのフィードバック刺激など、Oculofeedback 実験の実施に有用な情報までも、一括して出力できるようなアルゴリズムに改良し、眼球運動に関する情報処理を一手に担うようなものに発展させたいと考えている。

このようなシステムは、実験室の外における眼球運動の測定や解析を容易にすることから、眼球運動に関する研究のみならず、広範な応用研究にとって有用なシステムであると言える。

6.2 Oculofeedback の今後の展望

ヒトの脳のような複雑なシステムを理解するには、そのシステムでどのような信号がやり取りされているか、物理的にどのように実装されているかなど、複数のレベルで考える必要があると言われている [82]。その中でも、知性を機械的に実装するには、学習がどのように実現されているかを理解することが重要であると言われている [83]。

近年、人工知能の研究が活発になったことで、情報処理における学習への理解が進んでいる。

神経科学の視点では、Oculofeedback は被験者の学習を理解する枠組みとして利用できる。Decoded Neurofeedback では、被験者の学習のメカニズムとして、脳内で強化学習のような情報処理が働いていると考えることが合理的であるとされていることから [25]、Oculofeedback においても、そのような働きに関する知見の発見が期待できる。

一方、情報工学の視点からは、Oculofeedback のフィードバックを生成するアルゴリズムとして強化学習を利用できる。例えば、Neurofeedback の研究において、被験者の注意力を改善する実験に Deep Q Learning を応用する試みがなされている [84]。

このように、Oculofeedback は眼球運動の運動学習を理解するために重要な技術であり、ヒトや機械の学習に関する研究が追い風となる技術である。

今後は、益々発展する神経科学や情報処理の研究分野において、学習に関する理解が進むと同時に、Oculofeedback を実用的な技術として発展させていきたい。

謝辞

平田豊先生には日々の研究を支えていただき，熱心にご指導いただいたことに感謝します．また，国内外での貴重な発表の機会を数多く与えていただいたことに深く感謝いたします．研究室のメンバーには，実験や解析に参加していただき，貴重なデータを得ることに協力して頂きました．楽しい研究生活を送ることが出来たことを感謝します．また，津田一郎先生，藤吉弘亘先生，山下隆義先生には忙しい中，副査として博士論文の審査を引き受けて頂き深く感謝いたします．

参考文献

- [1] David L Sparks. The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience*, Vol. 3, No. 12, pp. 952–964, 2002.
- [2] Andrey R. Nikolaev, Sebastian Pannasch, Junji Ito, and Artem V. Belopolsky. Eye movement-related brain activity during perceptual and cognitive processing. *Frontiers in Systems Neuroscience*, Vol. 8, No. 62, pp. 1–2, 2014.
- [3] Andrew T. Duchowski. A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, and Computers*, Vol. 34, No. 4, pp. 455–470, 2002.
- [4] Andrew Duchowski. *Eye Tracking Methodology*. Springer-Verlag London, 2007.
- [5] Robert Gabriel Lupu and Florina Ungureanu. A survey of eye tracking methods and applications. *Buletinul Institutului Politehnic din Iasi, Automatic Control and Computer Science Section*, Vol. 3, pp. 72–86, 2013.
- [6] Katarzyna Harezlak and Pawel Kasprowski. Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics*, Vol. 65, pp. 176–190, 2017.
- [7] 増田正. 特集 3 人間工学のための計測手法 第 4 部 : 生体電気現象その

- 他の計測と解析 (1) . 人間工学, Vol. 51, No. 6, pp. 400–405, 2015.
- [8] 広田雅和, 不二門尚. 眼科領域における video-oculography を用いた眼球運動解析の研究動向. 視覚の科学, Vol. 38, No. 3, pp. 76–79, 2017.
- [9] 平田豊. 眼球運動を指標としたドライバ状態検知技術の実用化. 知の拠点あいち重点研究プロジェクト, 2018, <http://www.astf-kha.jp/project/project1/>.
- [10] B. Massé, S. Ba, and R. Horaud. Tracking gaze and visual focus of attention of people involved in social interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 40, No. 11, pp. 2711–2724, 2018.
- [11] 古川茂人, 米家惇, Hsin-I LIAO. 眼から読み取る心の動き : Heart-Touching-AI のキー技術 (特集 NTT グループにおける AI の取り組み). NTT 技術ジャーナル, Vol. 28, No. 2, pp. 22–25, 2016.
- [12] Sarah Uzzaman and Steve Joordens. The eyes know what you are thinking: Eye movements as an objective measure of mind wandering. *Consciousness and Cognition*, Vol. 20, No. 4, pp. 1882 – 1886, 2011. From Dreams to Psychosis: A European Science Foundation Exploratory Workshop.
- [13] 西山潤平, 木下伸一, 平田豊. 前庭動眼反射による眠気予兆検出. 生体医学工学, Vol. 48, No. 1, pp. 1–10, 2010.
- [14] 平田豊. 眼球運動と瞳孔ゆらぎによる眠気予兆検出. 信学技報 HIP, Vol. 112, No. 283, pp. 83–87, 2012.
- [15] 日本電気株式会社. NEC、離れた場所から視線の方向をリアルタイムかつ高精度に検知できる「遠隔視線推定技術」を開発, 2016. https://jpn.nec.com/press/201612/20161216_01.html, Accessed: 2019-12-17.
- [16] Erich Schneider, Thomas Villgrattner, Johannes Vockeroth, Klaus Bartl, Stefan Kohlbecher, Stanislavs Bardins, Heinz Ulbrich, and

- Thomas Brandt. EyeSeeCam: An eye movement-driven head camera for the examination of natural visual exploration. *Annals of the New York Academy of Sciences*, Vol. 1164, No. 1, pp. 461–467, 2009.
- [17] Akara Supratak, Chao Wu, Hao Dong, Kai Sun, and Yike Guo. Survey on feature extraction and applications of biosignals. *Machine Learning for Health Informatics: State-of-the-Art and Future Challenges*, pp. 161–182, 2016.
 - [18] Dana L Frank, Lamees Khorshid, Jerome F Kiffer, Christine S Moravec, and Michael G McKee. Biofeedback in medicine: who, when, why and how? *Mental health in family medicine*, Vol. 7, No. 2, pp. 85–91, 2010.
 - [19] Donald Moss. Biofeedback, mind-body medicine, and the higher limits of human nature. pp. 1–20, 1999.
 - [20] Thomas Budzynski, Johann Stoyva, and Charles Adler. Feedback-induced muscle relaxation: Application to tension headache. *Journal of Behavior Therapy and Experimental Psychiatry*, p. 205211, 1970.
 - [21] 辻下守弘. 【医学系】リハビリテーション医療におけるバイオフィードバックの応用. バイオフィードバック研究, Vol. 45, No. 1, pp. 41–44, 2018.
 - [22] 日本脳卒中学会脳卒中ガイドライン委員会. 脳卒中治療ガイドライン 2015 追補 2019 対応. 協和企画, 2019.
 - [23] 榊原雅人, 佐藤譲, 竹内聡, Park Suin, 及川欧. バイオフィードバック/ニューロフィードバックの臨床応用. 心身医学, Vol. 59, No. 7, pp. 613–621, 2019.
 - [24] Vincent J Monastra, Steven Lynn, Michael Linden, Joel F Lubar, John Gruzelier, and Theodore J LaVaque. Electroencephalographic biofeedback in the treatment of Attention-Deficit/Hyperactivity Disorder. *Applied Psychophysiology and Biofeedback*, Vol. 30, No. 2,

- pp. 95–114, 2005.
- [25] Kazuhisa Shibata, Lisi Giuseppe, Cortese Aurelio, Takeo Watanabe, Yuka Sasaki, and Mitsuo Kawato. Toward a comprehensive understanding of the neural mechanisms of decoded neurofeedback. *NeuroImage*, Vol. 188, pp. 539–556, 2019.
 - [26] Kazuhisa Shibata, Takeo Watanabe, Yuka Sasaki, and Mitsuo Kawato. Perceptual learning incepted by decoded fmri neurofeedback without stimulus presentation. *Science*, Vol. 334, No. 6061, pp. 1413–1415, 2011.
 - [27] 西村千秋. バイオフィードバックに必要な工学エッセンス. バイオフィードバック研究, Vol. 42, No. 1, pp. 65–69, 2015.
 - [28] Andrew D Straw. Vision Egg: An open-source library for realtime visual stimulus generation. *Frontiers in Neuroinformatics*, Vol. 2, No. 4, pp. 1–10, 2008.
 - [29] Jonathan W. Peirce. PsychoPy—psychophysics software in python. *Journal of Neuroscience*, Vol. 162, pp. 8–13, 2007.
 - [30] Carl Hewitt, Peter Bishop, and Richard Steiger. A universal modular ACTOR formalism for artificial intelligence. *Proceedings of the 3rd International Joint Conference on Artificial Intelligence*, pp. 235–245, 1973.
 - [31] Michael S. Gazzaniga. *The Cognitive Neurosciences, Fourth Edition*. MIT Press, 2009.
 - [32] 松嶋藻乃, 田中真樹. 空間的注意, 2012. <http://bsd.neuroinf.jp/wiki/空間的注意>.
 - [33] Yehoshua Tsao and Nilli Lavie. Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 19, No. 1, pp. 131–139, 1993.
 - [34] Jon Driver. A selective review of selective attention research from

- the past century. *British Journal of Psychology*, Vol. 92, No. 1, pp. 53–78, 2001.
- [35] Gordon L Corbetta, Maurizio; Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, Vol. 3, No. 3, pp. 211–215, 2002.
 - [36] Anna E Ipata, Angela L Gee, Jacqueline Gottlieb, James W Bisley, and Michael E Goldberg. LIP responses to a popout stimulus are reduced if it is overtly. *Nature Neuroscience*, Vol. 9, No. 8, pp. 1071–1076, 2006.
 - [37] Yair Pinto, Andries R. van der Leij, Ilja G. Sligte, Victor A. F. Lamme, and Steven H. Scholte. Bottom-up and top-down attention are independent. *Journal of Vision*, Vol. 13, No. 3, pp. 1–14, 2013.
 - [38] Fumi Katsuki and Christos Constantinidis. Bottom-up and top-down attention: Different processes and overlapping neural systems. *The Neuroscientist*, Vol. 20, No. 5, pp. 509–521, 2013.
 - [39] Michael I. Posner. Orienting of attention. *Quarterly Journal of Experimental Psychology*, Vol. 32, No. 1, pp. 3–25, 1980.
 - [40] John M Findlay and Iain D Gilchrist. Visual selection, covert attention and eye movements. In *Active Vision: The Psychology of Looking and Seeing*, chapter 3, pp. 35–54. Oxford University Press, 2003.
 - [41] 河原純一郎, 横澤一彦. 注意 : 選択と統合. 統合的認知, 第 1 卷. 勁草書房, 2015.
 - [42] John Findlay and Robin Walker. Human saccadic eye movements, 2012. Scholarpedia, http://www.scholarpedia.org/article/Human_saccadic_eye_movements.
 - [43] Maurizio Corbetta, Erbil Akbudak, Thomas E Conturo, Abraham Z Snyder, John M Ollinger, Heather A Drury, Martin R Linenweber,

- Steven E Petersen, Marcus E Raichle, David C Van Essen, and Gordon L Shulman. A common network of functional areas for attention and eye movements. *Neuron*, Vol. 21, No. 4, pp. 761–773, 1998.
- [44] Min Zhao, Timothy M. Gersch, Brian S. Schnitzer, Barbara A. Doshier, and Eileen Kowler. Eye movements and attention: The role of pre-saccadic shifts of attention in perception, memory and the control of saccades. *Vision Research*, Vol. 74, pp. 40–60, 2012.
 - [45] Susana Martinez-Conde, Stephen L Macknik, and David H Hubel. The role of fixational eye movements in visual perception. *Nature Reviews Neurosciences*, Vol. 5, No. 3, pp. 229–240, 2004.
 - [46] Susana Martinez-Conde, Stephen L. Macknik, Xoana G. Troncoso, and David H. Hubel. Microsaccades: a neurophysiological analysis. *Trends in Neurosciences*, Vol. 32, No. 9, pp. 463–475, 2009.
 - [47] 稲垣圭一郎, 臼井支朗. マイクロサッカードを制御する神経機構について. 日本神経回路学会誌, Vol. 19, No. 3, pp. 135–144, 2012.
 - [48] Ziad M Hafed and James J Clark. Microsaccades as an overt measure of covert attention shifts. *Vision Research*, Vol. 42, No. 22, pp. 2533–2545, 2002.
 - [49] Ralf Engbert and Reinhold Kliegl. Microsaccades uncover the orientation of covert attention. *Vision Research*, Vol. 43, No. 9, pp. 1035–1045, 2003.
 - [50] Susann Meyberg, Petra Sinn, Ralf Engbert, and Werner Sommer. Revising the link between microsaccades and the spatial cueing of voluntary attention. *Vision Research*, Vol. 133, pp. 47–60, 2017.
 - [51] Susana Martinez-Conde, Jorge Otero-Millan, and Stephen L Macknik. The impact of microsaccades on vision towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, Vol. 14, pp. 83–

96, 2013.

- [52] Jorge Otero-Millan, Xoana G. Troncoso, Stephen L. Macknik, Ignacio Serrano-Pedraza, and Susana Martinez-Conde. Saccades and microsaccades during visual fixation, exploration, and search: Foundations for a common saccadic generator. *Journal of Vision*, Vol. 8, No. 14, pp. 1–18, 2008.
- [53] Simone Benedetto, Marco Pedrotti, and Bruce Bridgeman. Microsaccades and exploratory saccades in a naturalistic environment. *Journal of Eye Movement Research*, Vol. 4, No. 2, pp. 1–10, 2011.
- [54] Shuntaro Miki and Yutaka Hirata. Microsaccades generated during car driving. In *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2148–2151, 2013.
- [55] 平山貴士, 三木俊太郎, 小栗崇治, 平田豊. 実車運転中の道路状況とマイクロサッカード発生特性. 信学技報 NC, Vol. 114, No. 226, pp. 81–86, 2014.
- [56] Takashi Hirayama, Takaharu Oguri, Keiichiro Inagaki, and Yutaka Hirata. Visual attention and individual miniature saccadic eye movements generated during car driving. Program No. 237.10/AA2. 2014 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2014. Online.
- [57] Jochen Laubrock, Ralf Engbert, Martin Rolfs, and Reinhold Kliegl. Microsaccades are an index of covert attention. *Psychological Science*, Vol. 18, No. 4, pp. 364–366, 2007.
- [58] Jochen Laubrock, Reinhold Kliegl, Martin Rolfs, and Ralf Engbert. When do microsaccades follow spatial attention? *Attention, Perception, and Psychophysics*, Vol. 72, No. 3, pp. 683–694, 2010.
- [59] Reinhold Kliegl, Martin Rolfs, Jochen Laubrock, and Ralf Eng-

- bert. Microsaccadic modulation of response times in spatial attention tasks. *Psychological Research*, Vol. 73, pp. 136–146, 2009.
- [60] 坂下祐輔, 藤吉弘亘, 平田豊. 面像処理による 3 次元眼球運動計測. 実験力学 : journal of JSEM, Vol. 6, No. 3, pp. 236–243, 2006.
- [61] Jorge Otero-Millan, Jose L. Alba Castro, Stephen L. Macknik, and Susana Martinez-Conde. Unsupervised clustering method to detect microsaccades. *Journal of Vision*, Vol. 14, No. 2, pp. 1–17, 2014.
- [62] Andra Mihali, Bas van Opheusden, and Wei Ji Ma. Bayesian microsaccade detection. *Journal of Vision*, Vol. 17, No. 1, pp. 1–23, 2017.
- [63] Ralf Engbert. Microsaccades: a microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, Vol. 154, pp. 177–192, 2006.
- [64] Mario Bettenbühl, Claudia Paladini, Konstantin Mergenthaler, Reinhold Kliegl, Ralf Engbert, and Matthias Holschneider. Microsaccade characterization using the continuous wavelet transform and principal component analysis. *Journal of Eye Movement Research*, Vol. 3, No. 5, pp. 1–14, 2010.
- [65] Shlomit Yuval-Greenberg, Elisha P. Merriam, and David J. Heeger. Spontaneous microsaccades reflect shifts in covert attention. *Journal of Neuroscience*, Vol. 34, No. 41, pp. 13693–13700, 2014.
- [66] Chih-Yang Chen and Ziad M. Hafed. Postmicrosaccadic enhancement of slow eye movements. *Journal of Neuroscience*, Vol. 33, No. 12, pp. 5375–5386, 2013.
- [67] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 8, pp. 1798–1828, 2013.
- [68] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning.

- Nature*, Vol. 521, No. 7553, pp. 436–444, 2015.
- [69] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network. *Computing Research Repository*, pp. 1–10, 2013. <https://arxiv.org/abs/1312.4400>.
 - [70] Yoshua Bengio. Practical recommendations for gradient-based training of deep architectures. *Neural Networks: Tricks of the Trade: Second Edition*, pp. 437–478, 2012.
 - [71] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *The International Conference on Learning Representations (ICLR)*, pp. 1–15, 2015.
 - [72] Masayuki Watanabe, Yuka Matsuo, Ling Zha, Douglas P. Munoz, and Yasushi Kobayashi. Fixational saccades reflect volitional action preparation. *Journal of Neurophysiology*, Vol. 110, No. 2, pp. 522–535, 2013.
 - [73] Makoto Yoneya, Shigeto Furukawa, and Makio Kashino. Potential use of microsaccade in personal identification. 37th annual international conference of the IEEE Engineering in Medicine and Biology Society, 2015.
 - [74] Yuk-Hoi Yiu, Moustafa Aboulatta, Theresa Raiser, Leoni Ophey, Virginia L. Flanagan, Peter zu Eulenburg, and Seyed-Ahmad Ahmadi. DeepVOG: Open-source pupil segmentation and gaze estimation in neuroscience using deep learning. *Journal of Neuroscience Methods*, Vol. 324, No. 108307, pp. 1–12, 2019.
 - [75] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen. MobileNetV2: Inverted residuals and linear bottlenecks. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, 2018.
 - [76] Alexander Mathis and Richard A. Warren. On the inference speed

- and video-compression robustness of DeepLabCut. *bioRxiv*, pp. 1–10, 2018.
- [77] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778. IEEE Computer Society, 2016.
 - [78] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. ImageNet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, Vol. 115, No. 3, pp. 211–252, 2015.
 - [79] Vincent Dumoulin and Francesco Visin. A guide to convolution arithmetic for deep learning. pp. 1–31, 2016. <https://arxiv.org/abs/1312.4400>.
 - [80] Eldar Insafutdinov, Leonid Pishchulin, Bjoern Andres, Mykhaylo Andriluka, and Bernt Schiele. DeeperCut: A deeper, stronger, and faster multi-person pose estimation model. In *European Conference on Computer Vision*, pp. 5–9, 2016.
 - [81] Aiga Suzuki, Hidenori Sakanashi, Shoji Kido, and Hayaru Shouno. Feature representation analysis of deep convolutional neural network using two-stage feature transfer-an application for diffuse lung disease classification. *IPSJ Online Transactions*, Vol. 11, No. 3, pp. 74–83, 2018.
 - [82] David Marr. *Vision*. MIT Press, 2010.
 - [83] Tomaso Poggio. The levels of understanding framework, revised. *Perception*, Vol. 41, No. 9, pp. 1017–1023, 2012.
 - [84] Martin Dinov and Robert Leech. Tracking and optimizing human performance using deep reinforcement learning in closed-loop

behavioral- and neuro- feedback: a proof of concept. *bioRxiv*, 2017.

研究業績

査読付き論文

- [1] 惠本 序珠亜, 平田 豊. 畳み込みニューラルネットワークによる実時間マイクロサッカード検出. 電子情報通信学会論文誌 D, Vol.J101-D, No.2, pp.456-467, 2018
- [2] Joshua Emoto, Yutaka Hirata. Lightweight convolutional neural network for image processing method for gaze estimation and eye movement event detection. Information Processing Society of Japan Transactions on Bioinformatics (TBIO), Vol.13, pp.1-9, 2020 (in print)

国際学会発表

- [1] Joshua Emoto, Yutaka Hirata. Development of oculo-feedback system: Manipulation of microsaccade generation. Society for Neuroscience, 531.18(Poster No.), San Diego, 2016
- [2] Joshua Emoto, Yutaka Hirata. Microsaccades predict spatial attention: Feature learning of microsaccade properties for oculo-feedback. Society for Neuroscience, 145.26(Poster No.), San Diego, 2018

国内学会発表

- [1] 惠本 序珠亜, 平田 豊. Oculo-feedback system の開発. 第 11 回空間認知と運動制御研究会, 2016.
- [2] 惠本 序珠亜, 平田 豊. 空間的注意向上のためのマイクロサッカードフィードバック実験. 第 31 回生体・生理工学シンポジウム 論文集, p.29,

2016.

- [3] 惠本 序珠亜, 平田 豊. Oculofeedback システムの開発とマイクロサッカードフィードバック実験. 信学技報 NC, 第 116 巻, pp.207-210, 2016.
- [4] 惠本 序珠亜, 平田 豊. 畳み込みニューラルネットワークを用いた微小眼球運動の検出. 人工知能学会全国大会 (第 30 回), 4L1-3, 2016.
- [5] 惠本 序珠亜, 平田 豊. Oculofeedback による注意向上訓練. 第 15 回注意と認知研究会, p.33, 2017.
- [6] 惠本 序珠亜, 平田 豊. 畳み込みニューラルネットワークを用いたマイクロサッカード検出法. 第 15 回情報学ワークショップ, p.40, 2017.
- [7] 惠本 序珠亜, 平田 豊. マイクロサッカード発生特性に基づく注意状態推定と眼球運動フィードバック実験. 日本視覚学会 2018 年冬季大会 抄録集, p.56, 2018.
- [8] 惠本 序珠亜, 平田 豊. マイクロサッカードによる空間的注意状態推定方法の検討. 第 13 回 空間認知と運動制御研究会 抄録集, p.15, 2018.
- [9] 惠本 序珠亜, 平田 豊. Oculo-feedback のためのマイクロサッカードと空間的注意の関係解析. 信学技報 NC, 第 113 巻, p.7, 2018.
- [10] 佐藤 健治, 吉田 傑, 石田 肇, 平田 豊, 惠本 序珠亜, 大前 学, 安部 原也, 内田 信行. 実車走行環境下における自動運転時のドライバ状態 (覚醒度) 検知に関する研究. 2018 年春季大会 自動車技術会 予稿集, No.139, pp.1-5, 2018

表彰

- [1] 平成 26 年度 中部大学, 学長表彰, 2015
- [2] 2016 年度 人工知能学会 全国大会 学生奨励賞, 2016
- [3] 平成 27 年度 電子情報通信学会 東海支部 奨励賞, 2016

- [4] 平成 28 年度 中部大学 学長表彰, 2017
- [5] 未来エネルギー研究協会 第 17 回 若手研究者のためのサマースクール,
ポスターセッション特別賞, 2017