

(別紙1)

## 総括研究報告書

課題番号	2023B-6						
研究開発課題名	国内で出生した乳児の顔写真収集と医療への活用を目指した社会基盤およびシステム基盤整備						
分類*	<input checked="" type="checkbox"/> ①	<input checked="" type="checkbox"/> ②	<input type="checkbox"/> ③	<input type="checkbox"/> ④	<input type="checkbox"/> ⑤	<input type="checkbox"/> ⑥	<input type="checkbox"/> ⑦
区分	<input type="checkbox"/> A	<input checked="" type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> E	<input type="checkbox"/> S		
主任研究者	所属	システム発生・再生医学研究部					
	役職	室長					
	氏名	岡村 浩司					
実施期間	2023年4月1日～2024年3月31日						

※分類は下記①～⑦より選択

- ① 日本の成育分野の疾患の研究の基盤となる研究
- ② 診断、治療及び予防法の開発に関する研究
- ③ 発症機序や病態の解明等を行う研究
- ④ 診断や治療のための基準の開発等に関する研究
- ⑤ 患児・者のQOL向上に結びつく研究
- ⑥ 研究的視点や技術をもつ医療従事者を育てるための研究  
(プロトコル作成のフェージビリティ研究)
- ⑦ 政策提言に結びつく研究

### 成果の概要

血縁関係がないにも関わらず顔が似ている人の遺伝子を比べると、塩基配列の類似性が高く、学歴など生き方まで似ているとの報告がある。その真偽はともかく、顔の学術的な価値は遺伝性疾患では以前から知られており、Noonan 症候群など、経験ある医師ならばその他多くの先天性疾患を指摘できるとされている。ディープラーニングによる画像認識の精度向上によりいくつもの成果が報告されており、国内でも人種の偏りを補正する独自データの必要性が主張されるものの、プライバシーの問題に阻まれ、作業は容易ではない。網膜芽細胞腫などの早期発見など、顔写真の有用性は眼科での報告も相次いでいる。小児の顔画像の医学的な価値は急速に高まる上に、AIにより他の分野への応用にも期待できる状況にある。成長により数年で変わってしまうものの、写真として残しておけば貴重な試料であり続ける顔写真のデータを、個人情報の保護を明確にした上で当センターがイニシアチブを取り、収集と医療への活用を目指した社会基盤およびシステム基盤整備を行う必要があると考えている。

眼科と協力して収集したデータ、140名の乳児から取得した763画像、およびFlickr-Faces-HQの公開データセットに登録されている顔写真から乳幼児の写真を1203枚探し出し、これらを組み合わせて研究を進めた。ディープラーニングを行うためには、数万程度のデータが必要であると言われていたが、特に希少疾患のサンプルを多数取得することは現実には不可能である。どの程度のデータ量で、どの程度の結果を出せるのか把握するため、明確に判別が可能な左右眼の判別AIモデルを、訓練数を制限して作成し、比較した。眼科データ763画像および公開1203画像に対し、手作業によるアノテーション作業を行い、それぞれ訓練データ、テストデータとして、検証を行った。1画像から左右眼を取り出すので、クロップ数は画像数の倍になる。ImageNetの転移学習を、フレームワークTensorflowを用いて実行し、認識精度0.9875、F値0.9873のきわめて良好なAIモデル作成を行うことができた。一方、クロップ数を制限した5モデルも作成し、ROC曲線を描いて調べると、クロップ数が100以上になるとF値は0.95を超え、さらに40クロップでも全クロップと同程度の精度を出すことができることを確認した。