

令和2年度放射線対策委託費
(放射線安全規制研究戦略的推進事業費)
放射線安全規制研究推進事業

染色体線量評価のための
AI自動画像判定アルゴリズム(基本モデル)の開発

国立研究開発法人量子科学技術研究開発機構
数藤由美子 (主任研究者)

1

課題名: 染色体線量評価のためのAI自動画像判定アルゴリズム(基本モデル)の開発

研究期間: 令和2年~令和3年(2年間)

背景・目的:

- 染色体異常を指標とする被ばく線量評価において、染色体異常の画像判定の標準化および効率化を目指して、人工知能(AI)技術のひとつである深層学習法を基盤とした染色体画像自動判定モデルの開発を行う。
- 本事業で開発するモデルは、量子科学技術研究開発機構の基本モデルとする。令和3年度には他機関での使用を考慮した汎用化に向けて、ギムザ染色画像や多様な品質の画像に対応するための技術検討やアプリケーション化のための技術検討を行う。

実施状況:

- 精度の高い線量評価の指標となる染色体異常の条件を明らかにした。
 - ・ 1回目分裂細胞の二動原体染色体
 - ・ 1回目分裂細胞の染色体断片 (世界で初めて証明 → Y. Suto et al., 2021, in press)
- PNA-FISH教師画像を作成中(1月末で13,197枚)。
- 同一標本でギムザ染色とPNA-FISHを行う技術、および同一細胞のギムザ画像とFISH画像を検出するための画像マッチング技術を確立した。
ギムザ画像とFISH画像を作成中(1月末で1,510組)。
- AIのアルゴリズムを検討・開発中(各種評価値が向上)。

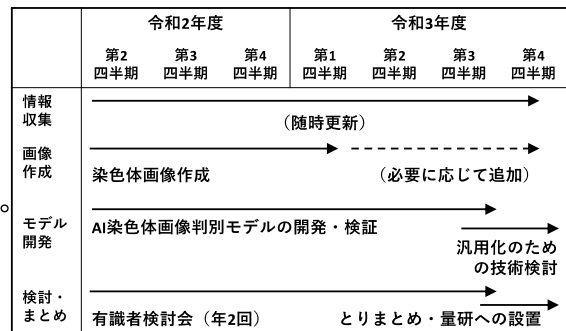


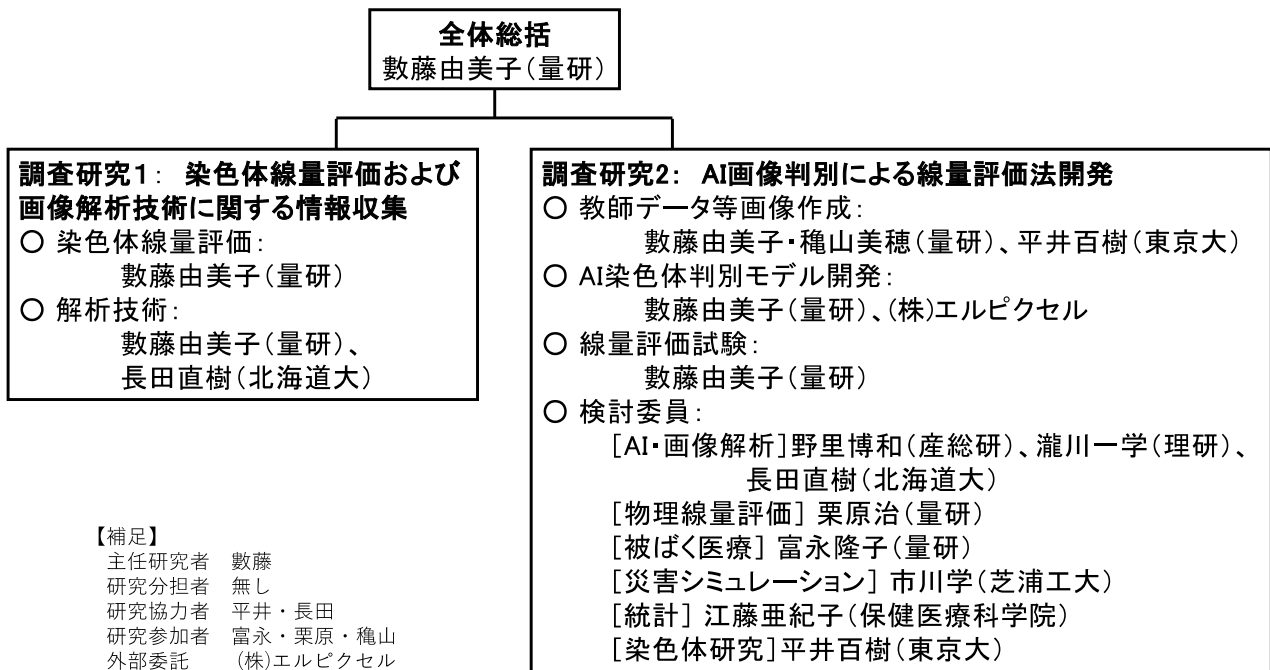
図1. ロードマップ

期待される成果:

- AIの導入により、画像判定が1検体につき約10分で可能となることが期待できる(コンピュータ並列化等によりさらに短縮)。大規模放射線事故での多検体トリアージ診断支援の大きな力となる。

2

研究実施体制



3

研究概要(1) 背景と目的

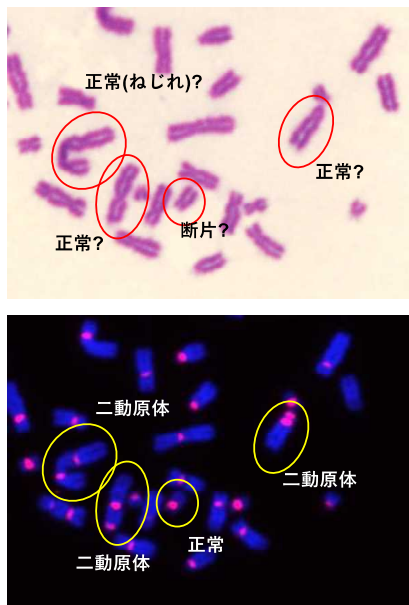
染色体線量評価では、あらかじめ線量と染色体異常頻度から成る検量線を作成し、患者の末梢血リンパ球の染色体異常頻度を当てはめて被ばく線量を推定する。染色体画像の作成までは装置の自動化により高速化されたが、画像判定は依然、検査者の目視観察による(患者1人当たりの判定に高度熟練者で実質30時間以上かかる)。熟練者養成は困難で、また、検査者により判定基準にブレが生じている。大規模原子力災害等に備え、染色体異常の画像判定の標準化と効率化が最大で喫緊の課題である。

一方、近年、分子細胞遺伝学の発展により染色体の染色方法の選択肢が増えた。並行して人工知能(以下、AI)による画像識別手法の開発には目覚ましいものがある。主任研究者らは平成30・31年度本研究事業において、染色体画像判定におけるAI導入の有効性の検証を行い、この分野での先鞭をつけた(以下、フィージビリティスタディ)。染色体線量評価において、PNA-FISH法を適用することにより、AI利用による染色体画像判定アルゴリズム開発が非常に有望であることが示された。(スライド5、参考図1・2参照)

以上の背景をふまえ、本研究では2年間で、AI技術のひとつである深層学習法を基盤とした染色体画像自動判定モデルの開発を目標とする。本事業で開発するモデルは、量子科学技術研究開発機構の基本モデルとし設置する。これによりトリアージのための線量評価の画像判定が1検体につき約10分以下で可能となる。加えて2年目後半には、汎用化に向けて、様々な品質の画像に対応するための技術検討、および、他機関での利用を容易にするギムザ染色画像の適用とアプリケーション化のための技術検討を行う。

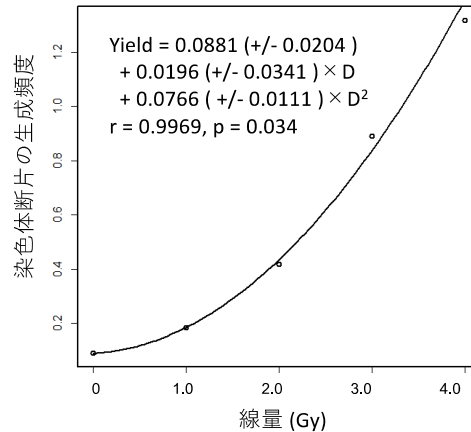
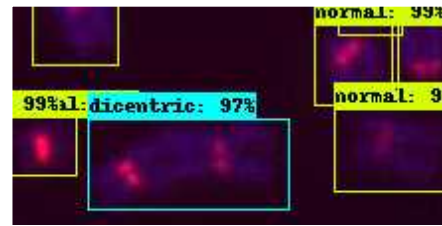
4

研究概要(1) 背景と目的 参考図



参考図1. 同一メタフェーズの部分図

ギムザ染色像(上)とPNA-FISH像(下)。PNA-FISHを用いれば、ギムザ染色より正確に染色体異常の判定ができるため、正確な教師画像が作成できることを示している。



参考図2. AI自動画像判定によりメタフェーズ(上)(部分表示)から検出された染色体断片を指標とした検量線(下) dicentric: 二動原体; normal: 正常; %表示: 確度。染色体断片を指標にすることができれば、二動原体染色体を指標とした場合より精度が高い可能性が示された。

「染色体線量評価手法の標準化に向けた画像解析技術に関する調査研究」令和元年度報告書より改変

5

研究概要(2) 令和2年度実施事項

(1) 線量推定のための指標としての染色体断片の評価(事業計画書2/4)

* フィージビリティスタディにおいて、染色体断片の有用性が示唆された。AI利用にあたって、線量推定理論への適合性と利用条件を明らかにした。

[材料] ⁶⁰Co-ガンマ線2.0 Gy照射末梢血リンパ球(線量率0.5 Gy/min)

[方法] 48~72時間培養 → [Fluorescent plus Giemsa (FPG)染色改変法]を併用した FISH解析から、細胞分裂回数ごとの染色体異常生成頻度を調べた。

(2) 教師画像データ作成(同 2/4)

* フィージビリティスタディにより、教師データの量と質が重要であることが示された。今年度、PNA-FISHの教師画像を作成した(増やした)。

* 汎用化検討に備え、同一細胞のPNA-FISH画像・ギムザ染色画像作成技術を開発し、作成を開始した。

(3) AIアルゴリズムの開発と検証(同 2/4)

* 2020年7月のAI学習モデルと人の目視観察の成績を比較した。

* フィージビリティスタディに基づき、アルゴリズムのさらなる検討・開発を進めた。

(4) 情報の更新(同 1/4)、成果の公表(同 3/4)、進捗管理(同 4/4)

* 本研究の推進にあたり、適宜、AI・画像解析技術・染色体線量評価法に関する情報収集、成果発表、進捗報告(月末報告書提出)を行っている。

6

研究の進捗(1)

[Y. Suto et al., 2021, in press]

(1) 線量推定のための指標としての染色体断片の評価

- 1細胞当たりの染色体断片生成数の分布はポアソン分布にしたがうか？
 - ① カイ二乗適合度検定
 - ② 分散比 (1,000細胞を解析)

- 第1回目分裂細胞

- ① $0.10 < p < 0.25$
- ② 1.05 (ポアソン分布にしたがう)

- 第2回目分裂細胞

- ① $p < 0.05$
- ② 1.72 (過分散)

→ 染色体断片は二動原体(図2-a)だけでなく、欠失(図2-b)や環状染色体(図2-c)の生成等でも生じる。また、第2回目分裂細胞では、染色体不分離によりツインの染色体断片が生じたり(図2-a, b)、染色体断片が非同調複製/凝縮を起こしてイレギュラーな形態をもったり(図2-c)することがわかった。それらが原因の一部であると考えられる。

【結論】

* 第1回目分裂細胞の染色体断片は線量推定のための指標に使える。本研究で本格的に導入する(二動原体と併用する)。

→ 二動原体と違い、染色体断片は染色体技術の格差の影響が抑えられるので、汎用化に有用。

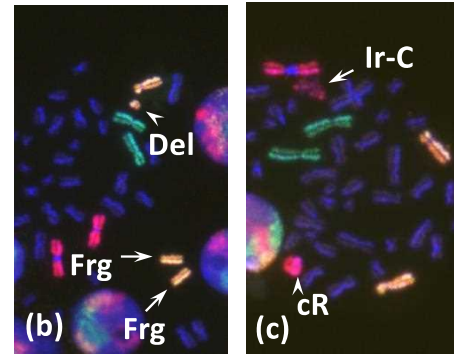
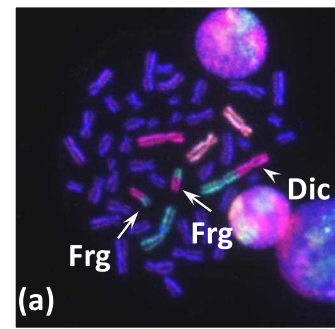


図2. 第2回目分裂細胞の3-color FISH像
赤: 1番染色体; 緑: 2番染色体; 黄: 4番染色体; 青: 対比染色。Dic: 二動原体; Frg: 染色体断片; Del: 欠失; cR: 環状染色体; Ir-C: イレギュラーな形をとったクロマチン。

(参考) 二動原体染色体の場合、第1回目分裂細胞: ① $0.50 < p < 0.75$ ② 1.02; 第2回目分裂細胞: ① $0.25 < p < 0.50$ ② 1.07

7

研究の進捗(2)

(2) 教師画像データ作成

* 画像検討会10回実施 (2月、3月分予定を含む)

- ^{60}Co -ガンマ線0~4.0 Gy照射末梢血リンパ球染色体標本を用いてメタフェーズ教師画像を作成中

1) PNA-FISHの教師画像作成

* 13,197枚の画像を準備 → 自動アノテーション・目視修正を実行中(現在 2,700枚: 0 Gy (12%), 0.50 Gy (14%), 1.0 Gy (16%), 2.0 Gy (32%), 3.0 Gy (17%), 4.0 Gy (9%); 二動原体染色体数1,080個、染色体断片数1,849個で、昨年度比各1.9倍)。目視で高品質と判断した画像割合は0~3.0 Gyで平均 0.80 ± 0.05 、4.0 Gyで0.59であった。

2) 同一細胞のPNA-FISH画像・ギムザ画像作成 (特許申請のため非公開) (現在 1,510組)

(3) AIアルゴリズムの開発と検証

- 人の目視判定結果とAI判定結果の成績比較

^{60}Co -ガンマ線3.0 Gy照射末梢血由来染色体標本のPNA-FISH画像200枚を用い、人5名とAI(2020年7月版モデル2種*)で判定成績を比較した(図3)。さらにそれぞれの検量線を用いて線量推定を行った。

【人】 $2.6 \pm 0.24 \sim 3.1 \pm 0.24$ Gy (二動原体)

【AI】 2.5 ± 0.43 Gy (二動原体)

3.0 ± 0.27 Gy (染色体断片)

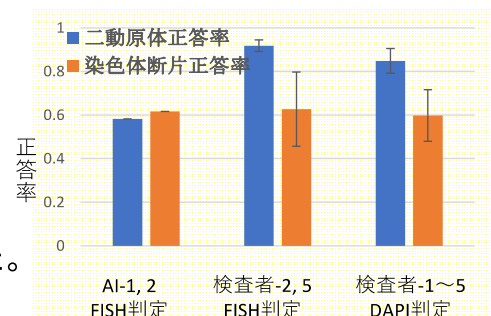


図3. AI(2種)および人(5名)の染色体異常検出正答率
人ではギムザ染色の代替としてDAPI画像の観察も行った。

* フィージビリティスタディ試作版と同様のFaster-RCNNを用いたモデルで、染色体を正常・二動原体・断片・その他の4つに分類して試作したものと、全ての染色体異常を区別し分類して試作したものの2種を用いた。

8

研究の進捗(3)

(3) AIアルゴリズムの開発と検証 (つづき)

* 研究会合2回実施 (2月、3月分予定を含む)

● アルゴリズム開発(進行中)

フィージビリティスタディで使用したFaster-RCNN以降に登場した3つの手法を加え、検討と改良を進め、全体的に向上しつつある。ここでは**モデルA [Faster-RCNNの改良型(2020年12月版)]**、B、C、Dのうち、平均絶対誤差が良好と判明した、昨年度フィージビリティスタディ・試作版の更新版であるAと、Bを例に挙げる。

- 1) 平均絶対誤差 (AIが自動検出した1細胞中の染色体数と真の値との差の絶対値)
→ Bで誤差が半減した(図4)。
- 2) 精度 (Precision; 陽性の信頼度) と感度 (Recall; 陽性を見逃さない割合) (図5)
→ 二動原体検出の感度が大きく向上、染色体断片でも向上した。
- 3) AI自動判定による線量推定試験(70メタフェーズ×3組)(図6)
→ 染色体断片はA・B同傾向で良好。二動原体はBで改善された。

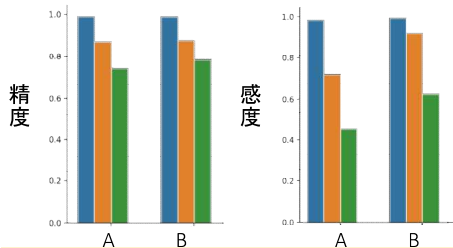
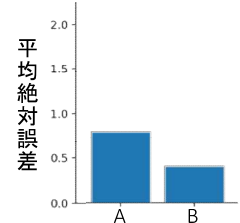


図5. モデルAおよびBの精度と感度の比較
■ 正常染色体; ■ 染色体断片; ■ 二動原体

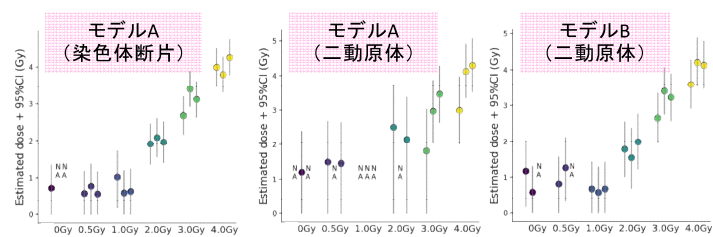


図6. モデルAおよびBによる線量推定試験(スライド10の補足図参照)
メタフェーズ70個を1組とし、3組をテストした。NA: CI下限値が0以下で適用外。

研究の進捗(3) 図6の補足(拡大・補足図)

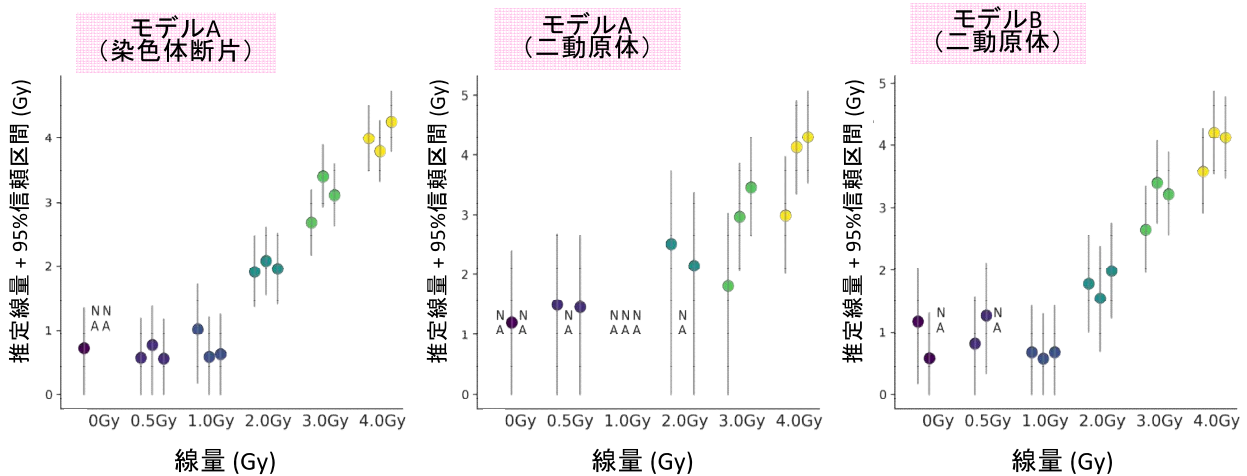


図6. モデルAおよびBによる線量推定試験

メタフェーズ70個を1組とし、3組をテストした。NA: CI下限値が0以下で適用外。

研究の進捗(4)

(4)情報の更新、成果の公表、進捗管理

● 情報収集

- ① 日本人類遺伝学, 第65回大会(ウェブ開催、令和2年11月18日～12月2日)
- ② Radiation Research Society, 66th annual meeting, Symposium 17 (Biodosimetry), USA (ウェブ開催、令和2年10月20日)
- ③ Incident and Emergency Centre, IAEA, Webinar: Biodosimetry: from fundamental science to practical application in response to nuclear and radiological safety or security related emergencies (令和2年10月15日).
- ④ ISO/TC85/SC2 (Radiological Protection) 会議; WG-18(被ばく事故発生時の二動原体分析手順書の策定)、WG-25会議(被ばく事故・住民等対応手順書の策定)(令和2年11月10日～11月14日)
- ⑤ 文献・ウェブ情報探索(米国グループによる二動原体解析自動化検討など)

(注) 参加予定していた他の研究会合(Biodosimetry関連1件、AI・画像解析関連2件)は新型コロナウイルスの影響により来年度に延期)

- 成果の公表 → スライド12参照(スライド8・9の成果は特許申請を検討中)
- 進捗管理 → 毎月末、POおよびPO補佐に報告書を提出

11

成果

【原著論文】

- 1) Yumiko Suto, Takako Tominaga, Miho Akiyama, Momoki Hirai: Revisiting Microscopic Observation of Chromosomal Aberrations in Cultured Human Peripheral Blood Lymphocytes at the Second Mitotic Division after Gamma Irradiation In Vitro. *Cytologia* 86(1), 2021, in press. (令和3年3月刊行予定)

【招待講演】 (本研究および前事業フェーズビリティスタディの両方の成果を含めて発表した)

- 1) Yumiko Suto: Cytogenetic biodosimetry of radiation accidents. Symposium 17: Biodosimetry, Radiation Research Society's 66th annual meeting. (2020年10月20日、ウェブ開催)
- 2) Yumiko Suto: Effective utilization of biodosimetry at Fukushima Daiichi Accident. Webinar: Biodosimetry: from fundamental science to practical application in response to nuclear and radiological safety or security related emergencies. (Incident and Emergency Centre, IAEA, 2020年10月15日、ウェブ開催).
- 3) 数藤由美子: 放射線被ばくによる染色体異常. 第27回 臨床細胞遺伝学セミナー.(日本人類遺伝学会、ウェブ開催、2020年10月16日～11月1日配信)

12

関連の成果

【国際的評価】

- ① 国際標準化機構ISO WG-18(生物線量評価ワーキンググループ)、2020年11月会議にて、本研究事業の予備研究である令和元年度放射線安全規制研究戦略的推進事業の成果論文を米国委員が採り上げ、議題となった。手順書 19238 および ISO 21243 の次回改訂において採用する方向で検討に入った(各国ラボで追試中)。

* Yumiko Suto, Takako Tominaga, Miho Akiyama, Momoki Hirai: Cytogenetic Examination of Human Peripheral Blood Lymphocytes Cryopreserved after Gamma Irradiation: A Pilot Study. Cytologia 85(1): 71-77, 2020.

- ② 図説提供(本研究事業関連の染色体技術による)

1) Genetik, 7th Ed. Springer, Germany(2021, in press, 3月刊行予定)
(ISBN 978-3-662-60909-5).

2) BBC Horizon, BBC, UK. (使用許可期間 2020年～2030年)

13

自己評価

1. 研究代表者(主任研究者)による自己評価

評価の視点	自己評価	コメント
評価時点までの研究の実施が研究計画に沿って行われているか	2. 概ね計画通り	1)フィージビリティスタディで示唆された染色体断片の有用性について、正式に染色体断片が線量評価の指標として使用できることを証明し、使用条件を明らかにした(Ⅱ 評価時までの研究成果、【原著論文】参照)。以上の結果をAI画像判定に適用することとした。 2)機械学習(深層学習)モデルを複数検討した。二動原体・染色体断片ともに検出性能が向上しつつある。来年度、染色体が正しく検出できていない事例の原因探索と染色体検出性能のモデル差の原因探索により、更に性能を向上させる。 3)PNA-FISHの教師画像作成を行った(13,197枚)。また、他機関への汎用化検討の準備として、同一細胞のPNA-FISH画像・ギムザ染色画像作成技術を確立した(1月末時点で1,510組作成)。来年度も引き続き教師画像を作成し、トレーニング量を増やすことでも、更にAI判定の性能を向上させる。
今年度の進捗や達成度を踏まえて、次年度の研究計画に変更が必要か	2. 軽微な変更が必要	令和3年度も引き続き研究会合は原則としてウェブ開催とする。そのために予定していた旅費は、教師画像作成のための画像データ検討会に充当し、可能な限り画像を増やす。

2. 分担研究者による自己評価 → (分担研究者 無し)

14