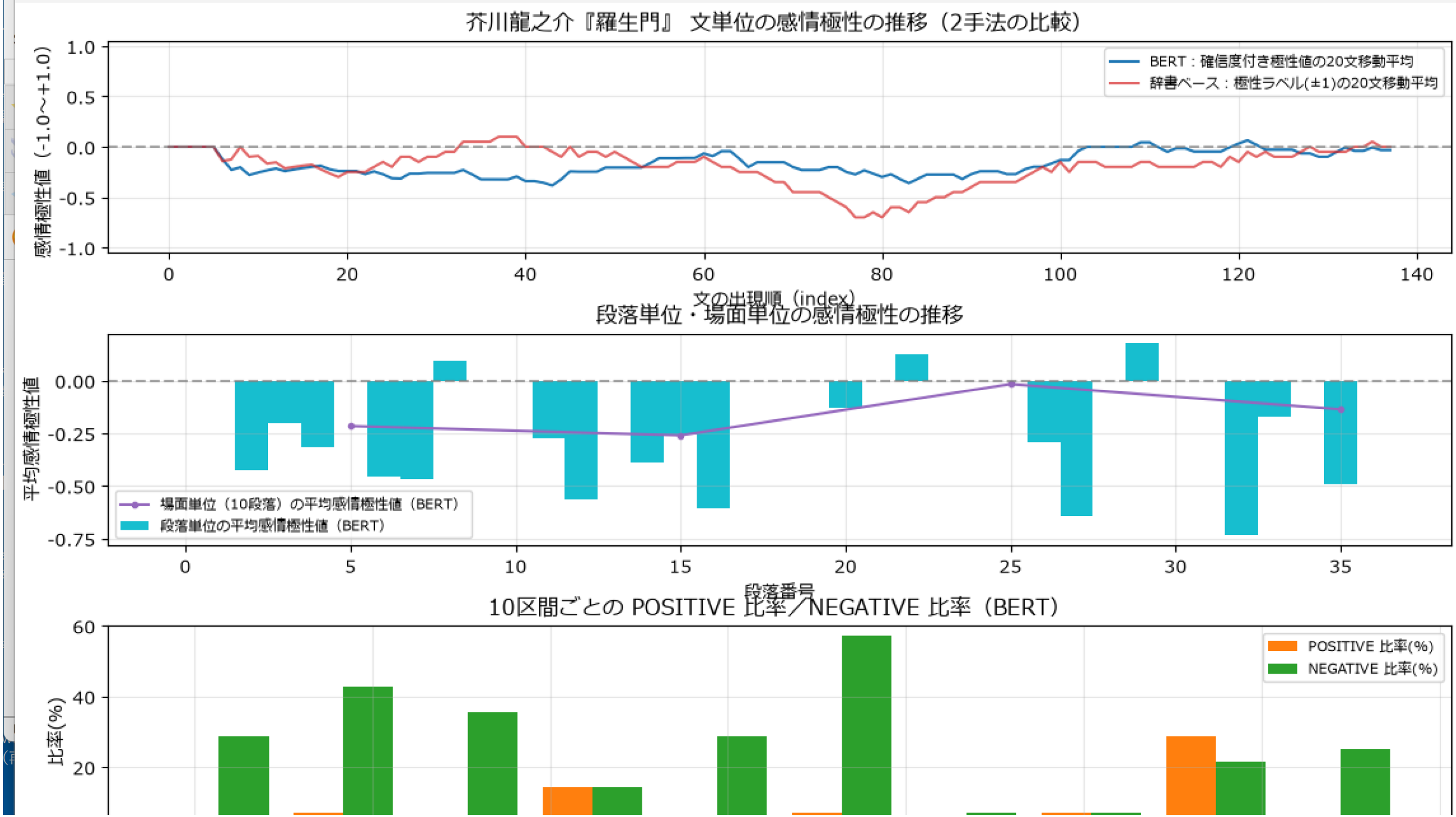
1.**日本語文学作品の感情分析**（学習済み二ューラルモデル（BERT）と日本語評価極性辞書の比較）

芥川龍之介『羅生門』の感情推移



日本語文学作品の感情分析（学習済みニューラルモデル（BERT）と各日本語評価極性辞書の比較実全体内の実験流れ

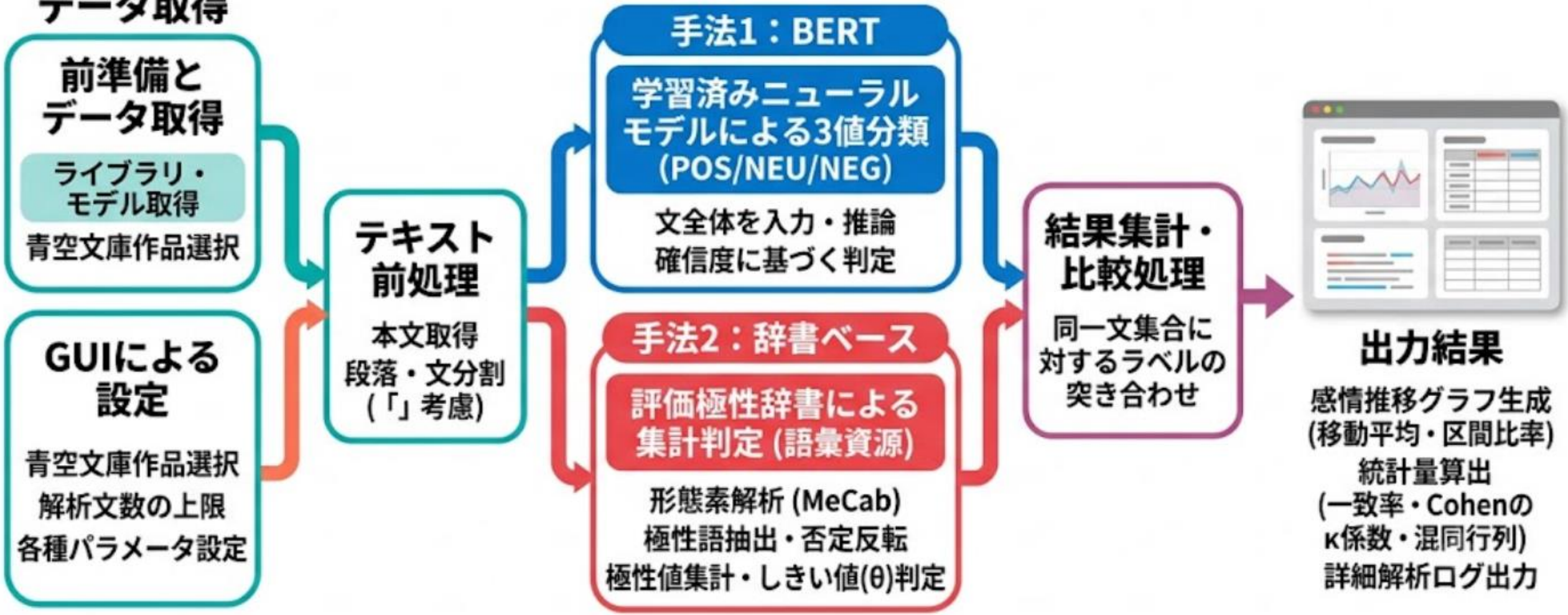
前準備とデータ取得

プログ処理

手法例分析

結合

出力結果



日本語文学感情分析 比較実験ツール

使用技術と手法の概要

手法1: 学習済みBERT（AIモデル）
koheiduck/bert-japanese-finetuned-sentiment

手法2: 日本語評価極性辞書（語彙資源）
東北大学 乾・岡崎研究室

実装技術
Python, Tkinter (GUI), fugashi, ipadic,
transformers, datasets, matplotlib

AIモデル（BERT）の特徴と学習済みデータ

model id:
koheiduck/bert-japanese-finetuned-sentiment
base model:
cl-tohoku/bert-base-japanese-whole-word-masking

事前学習データ
- 日本語Wikipedia (2019年9月1日時点のダンプ)

モデルの特徴
- 文全体を肯定・中立・否定に3値分類
- 語順・否定を含む文脈を考慮した判定
- トークナイザ: MeCab+IPAdic (fugashi, ipadic)
- 特記事項: 微調整データは非公表

日本語文学作品の感情分析 比較実験

1. 実験目的

青空文庫を
文単位で
3値分類
(POS/NEU/NEG)

原理の異なる
2手法の
整合性を
定量化

2. 対象データと手法

対象データ：
青空文庫クリーン
データセット

手法A：
学習済み
BERT
(koheiduck/...)
[AIモデル]

文脈を考慮、
確信度出力

手法B：
日本語評価
極性辞書
[語彙集計]

単語ベース、
最長一致、
否定反転

3. 提示必須の結果指標

- 一致率 (Po)
- Cohenのκ係数 (kappa)
- 3×3 混同行列
- 感情推移グラフ (移動平均)

4. 考察の最重要点

κ係数は「精度」ではなく「整合性」の指標
(正解ラベルなし)

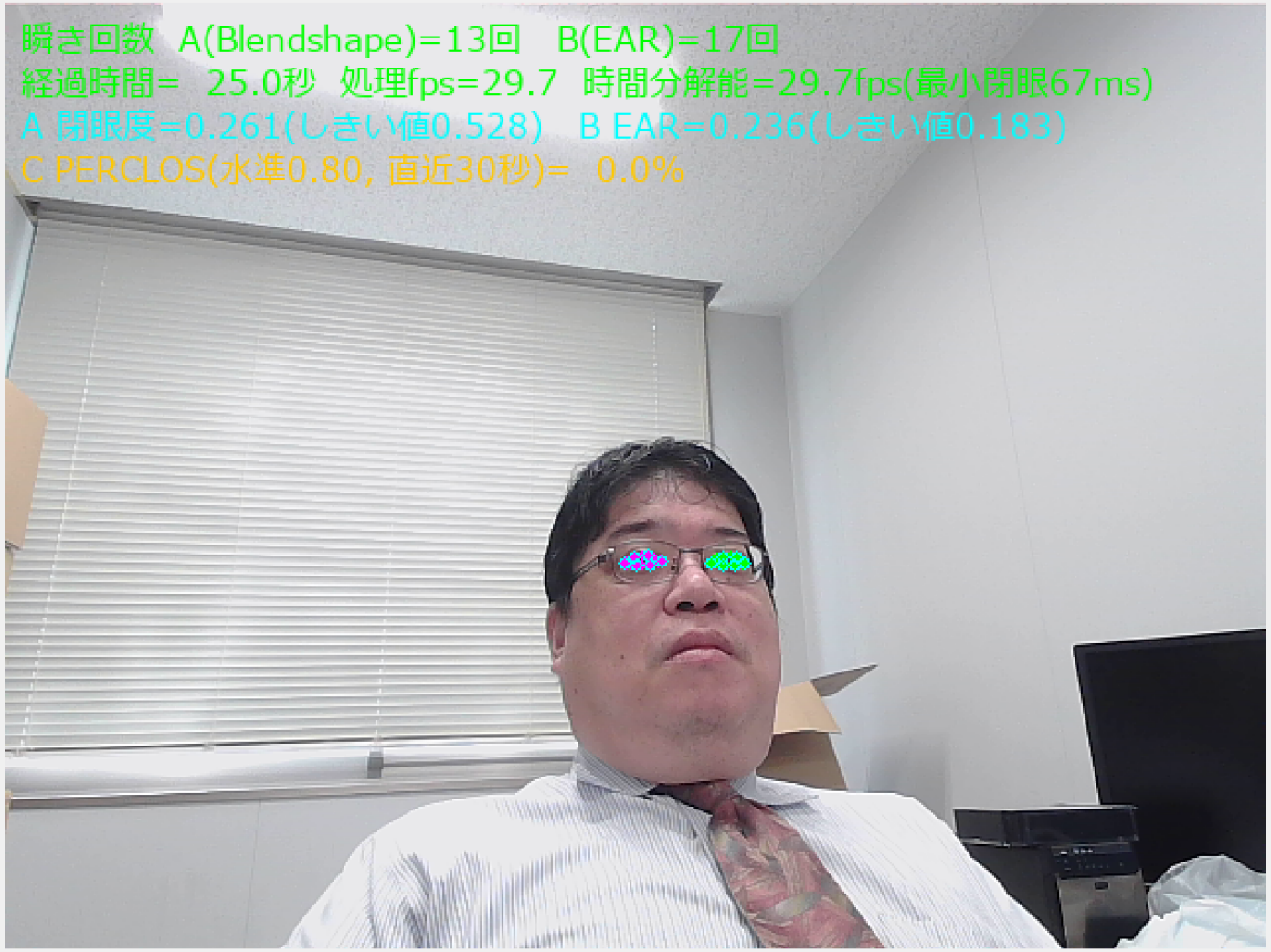
辞書ベース手法は
NEUTRALに偏る傾向
(一致率上昇要因)

近代文学の語彙・文体に対する現代モデル・辞書の領域のずれ

パラメータ設定
(しきい値θなど)
による結果の変動

2. 瞬目検出3手法の比較と瞬き頻度の時間推移の観察

瞬目検出3手法の比較（Blendshape / EAR / PERCLOS）



1. 入力の選択

カメラ　カメラ番号 0

動画ファイル　ファイルを選択

（未選択）

2. パラメータ

手法A 閉眼判定マージン	0.30
手法B EAR比率しきい値	0.75
最小閉眼持続フレーム数	2
基準値推定の窓長(フレーム)	60
手法C PERCLOS閉眼水準	0.80
手法C 移動窓長(秒)	30
集計区間長(分)	2

計測開始　計測停止　結果を集計して保存

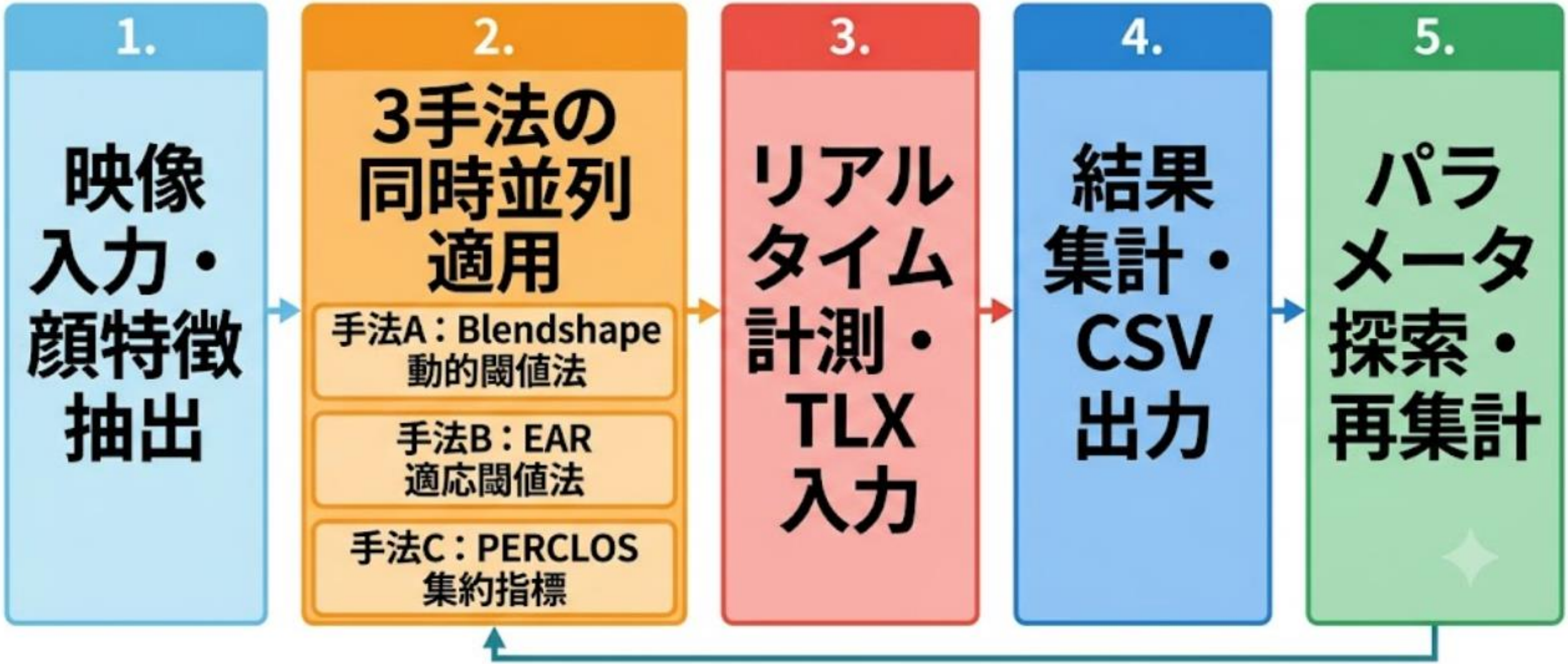
集計完了：blink_log.csv と signal_log.csv を保存した

3. NASA-TLX簡易版（カメラで自分を計測した回のみ、計測終了後に1回入力）

精神的要求	50	身体的要求	50
時間的切迫感	50	作業成績	50
努力	50	フラストレーション	50

```
================================================================
入力：カメラ
総フレーム数 736    顔検出失敗 0フレーム (0.0%)
計測時間 0.42分    時間分解能 29.5fps    最小閉眼持続 68ms
----------------------------------------------------------------
手法A Blendshape :    13回    平均  31.2 回/分
手法B EAR        :    17回    平均  40.8 回/分
手法C PERCLOS    : 平均    0.0%    最大    0.0%
AとBの対応 : 一致13回 / Aのみ0回 / Bのみ4回
一致度(Dice係数 2M/(Na+Nb))= 0.867    ±0.5秒以内を同一とみなす
  ※どちらも正解データではないため、これは一致の度合いであり精度ではない
----------------------------------------------------------------
区間別の瞬き頻度(区間長 約2分。端数は最終区間に含める)
    0.0〜  0.4分   A:  13回( 31.2/分)    B:  17回( 40.8/分)
```

「瞬目検出3手法の比較」実験全体の流れ



瞬目検出3手法の比較と瞬き頻度の時間推移の観察

MediaPipe Face Landmarker (AIモデル)

構成と特徴

BlazeFace (顔検出)
FaceMesh-V2 (478点ランドマーク)
Blendshape V2 (52表情係数)

MLP-Mixer系、146点から高速推定

学習済みデータ

多視点顔画像、3D GHUMメッシュ、
人物・表情・向きの合成データ(非公開)

手法A: Blendshape動的閾値法

- 入力：eyeBlink係数平均（0〜1）
- 閾値：履歴下位10%+マージン
- 出力：瞬き回数・時刻

手法B: EAR適応閾値法

- 入力：EAR平均（左右6点）
- 閾値：履歴上位90%x比率(既定0.75)
- 出力：瞬き回数・時刻

手法C: PERCLOS

- 入力：閉眼度
- 特徴：移動窓内の閉眼水準水準（既定0.80)以上割合
- 出力：閉眼時間率（%）

共通：A・Bは閉眼持続後開眼で計数

特徴：Cは時間割合、眠気指標

プログラムの出力と目的

出力

画面表示（回数、時間、FPS、閾値、PERCLOS）、CSVログ、
集計結果（総数、頻度、相関、一致度、Raw TLX）。

目的

3手法定量比較、時間推移観察、
パラメータ探求、主観負荷との傾向確認

瞬目検出3手法の同時比較 と 瞬き頻度の時間推移の観察

3つの検出手法（入力: MediaPipe）

A. Blendshape法 (eyeBlink係数)
（表情係数による閉眼度推定）
→目が後による閉眼を増す

B. EAR法 (Eye Aspect Ratio)
（目の幾何的形状（縦横比））
→目が幾による閉眼を減す

C. PERCLOS (Percent of Closure)
（閉眼時間率（眠気指標、回数外））

主要な出力・評価項目

瞬き回数（A vs B）

平均頻度（回/分）

手法間一致度
(Dice係数)

主観負荷 (Raw TLX)

実験フロー と 考察のポイント

【手順】
動作確認 → 負荷作業 → 集計

【考察】
正解データなし。手法間の定量比較。

パラメータ設定（閾値・窓長）の影響。

3.**静止画分類器と動画分類器による**
フェイク動画判定の比較（ViT / SigLIP / VideoMAE を同一動画に同時適用

静止画分類器と動画分類器によるフェイク動画判定の比較（教材用）

1. 入力動画の選択

カメラ　カメラ番号 0

動画ファイル　ファイルを選択　（未選択）

Hugging Face データセット　video/1.mp4　UniDataPro/deepfake-videos-dataset　（video/ = 元動画, deepfake/ = 加工動画）

2. パラメータ

静止画分類のサンプリング間隔［フレーム］ 10

顔クロップの余白率 0.35

偽側と判定するしきい値 0.50

クリップ構成のフレーム間引き幅 1

Farneback 法の窓サイズ［画素］ 15

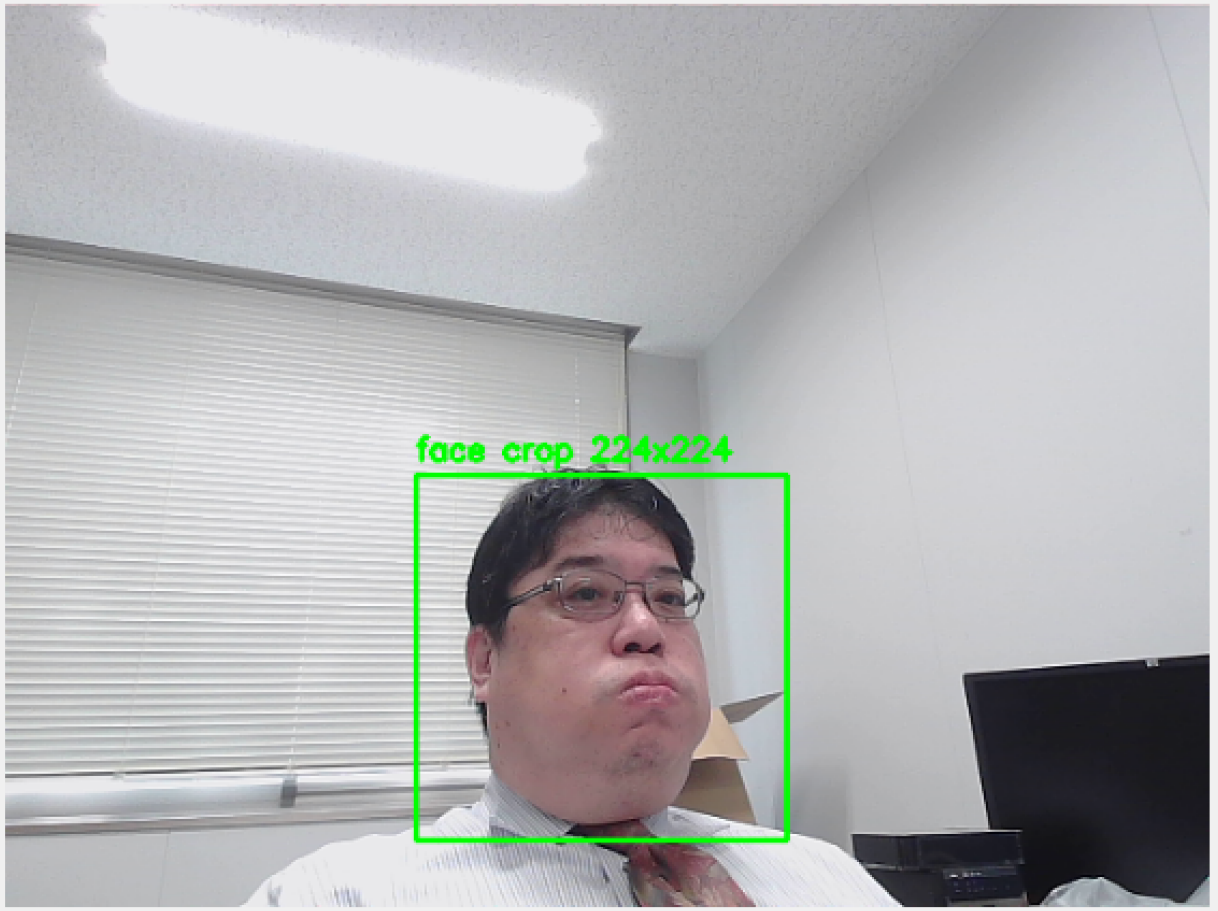
顔検出の最小信頼度（開始時に反映） 0.50

クリップ長 16 フレーム / 入力 224x224 画素はモデル仕様のため固定

3. 操作

解析開始（取り込み開始）　解析停止（取り込み停止）

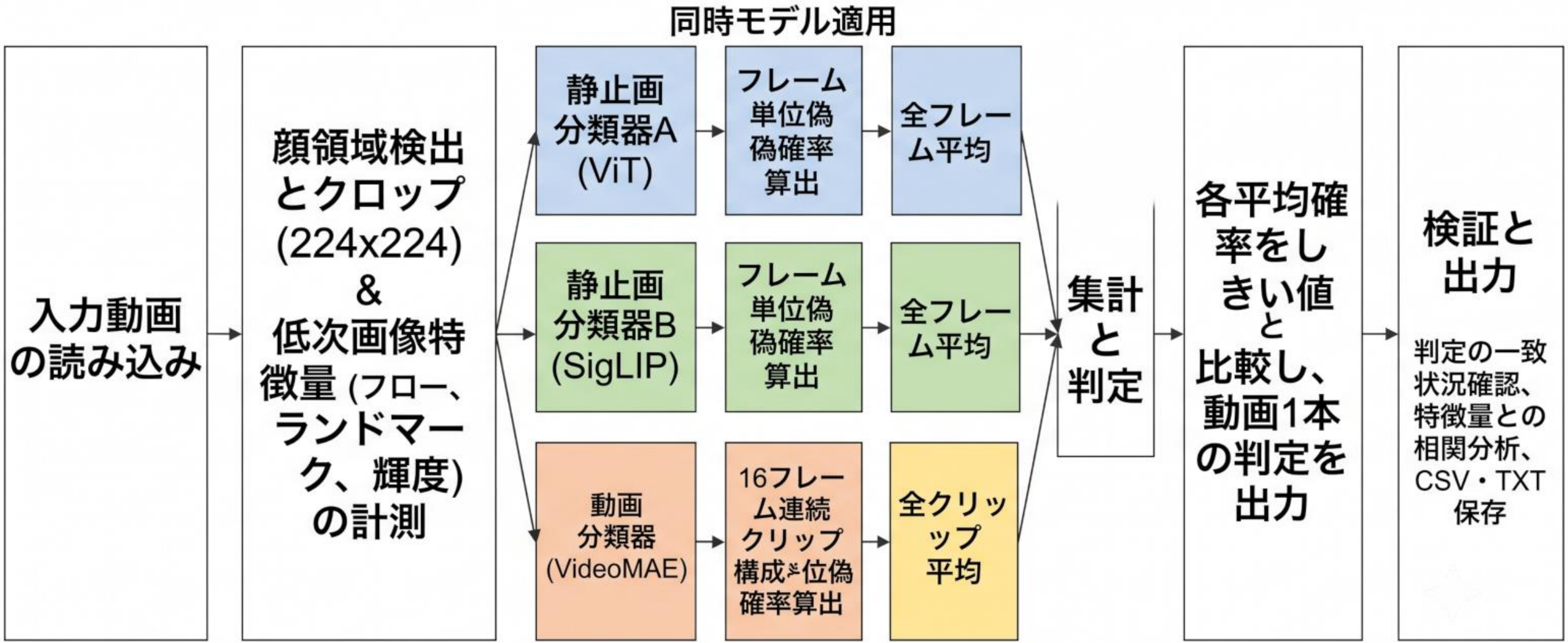
4. 入力映像と顔クロップ領域（緑枠）



5. 解析ログと最終判定

```
ViT の Deepfake 確率 と オプティカルフロー平均: r = -0.192 (標本数 30)
SigLIP の Fake 確率 と オプティカルフロー平均: r = -0.029 (標本数 30)
ViT の Deepfake 確率 と ランドマーク変位: r = +0.397 (標本数 30)
SigLIP の Fake 確率 と ランドマーク変位: r = +0.174 (標本数 30)
```

[実験全体の流れ：静止画・動画分類器によるフェイク判定比較]



フェイク動画判定モデル比較（探究用ガイド）

入力：
同一動画の
同一顔領域
クロップ
(224x224)

静止画分類器 A (ViT)

使用技術	AIモデル特徴	学習済みデータ
静止画1枚入力, Patch分割 Vision Transformer (ViT)	prithivMLmods/Deep-Fake-Detector-v2-Model google/vit-base-patch16-224-in21k 基盤	ImageNet-21k 事前学習 微調整は非公開

静止画分類器 B (SigLIP)

使用技術	AIモデル特徴	学習済みデータ
静止画1枚入力 視覚言語事前学習モデル	prithivMLmods/deepfake-detector-model-v1 SigLIP	微調整 prithivMLmods/ OpenDeepfake-Preview

動画分類器（VideoMAE）

使用技術	AIモデル特徴	学習済みデータ
連続16フレーム入力 Tubelet分割 時空間 Transformer	Ammar2k/videomae-base-finetuned-deepfake-subset MCG-NJU/videomae-base 基盤	Kinetics-400 事前学習 微調整は非公開

フェイク動画判定の比較

実験目的：比較の核心

- フェイク検出における「時間情報の効果」を検証
- 静止画分類器と動画分類器を同一対象に同時適用
- 動画レベルでの判定一致・不一致を突き合わせる

実装モデルと適用方法

- 入力：顔クロップ（224×224）のみ。音声不使用
- 静止画分類器（フレーム単位）：ViT / SigLIP
  フレームごとに確率を出力、動画全体で平均
- 動画分類器（クリップ単位）：VideoMAE
  連続16フレームをまとめて処理、動画全体で平均

検証数値と低次特徴量

- 補助指標（低次特徴量）：
  - オプティカルフロー平均
  - ランドマーク変位
  - 輝度差（左右・上下）
- 検証：分類器確率と低次特徴量の相関
  ※低次特徴量は分類器の入力ではない

考察と本教材の限界

- 各モデルの真の検出精度は不明（正解ラベル不足）
- VideoMAEの学習データ・精度は非公開
- 判定の内部根拠（顕著性マップ等）は出力不可
- 顔領域のみ対象。背景・全画面は扱わない

4.**画像からの異常検知**（統計モデル（PaDiM）・メモリバンク（PatchCore）・基盤モデル特徴（DINOv2最近傍法）の比較）

AI体験学習：画像からの異常検知（PaDiM / PatchCore / AnomalyDINO）

正常画像の枚数
20

PaDiM 使用特徴次元数
100

PatchCore サンプリング率 [%]
10

PatchCore 近傍数 k
1

異常判定しきい値（スコア比） [%]
150

正常画像の登録

処理停止

カメラ映像（ライブ）

PaDiM

スコア比 127 % 判定: 正常

PatchCore

スコア比 155 % 判定: 異常あり

AnomalyDINO

スコア比 339 % 判定: 異常あり

異常マップの色: 青=異常度低、赤=しきい値到達（基準値としきい値で正規化）

連続処理中です。被写体に異常を作って3手法の反応を比較してください。

NanoBanana2: 画像からの異常検知 3方式比較実験全体の流れ

Step 1: 実験ツールの準備

Step 1:
実験ツールの準備

Python、
PyTorch環境、
ネットワークモデル
(ResNet-18,
DINOv2)の取得

Step 2: 被写体とカメラの固定

Step 2:
被写体と
カメラの固定

静的な被写体を
カメラに向け、
照明を安定させる

正常登録

PaDiM (統計モデル)

正常特徴からパッチ位置ごとの多変量ガウス分布を推定

PatchCore (メモリバンク)

正常パッチ特徴のメモリバンクを構築 (コアセット選択)

DINOv2最近傍法 (基盤モデル特徴)

DINOv2パッチ特徴を参照メモリとして保持

Step 3: 正常状態の登録
(複数枚撮影・モデル構築)

Step 4: 連続検知と観察

実験の進め方 (変化を加える)

- 異物を置く
- 被写体を少し動かす
- 照明を変える

観察と考察ポイント (3方式を比較)

- 各方式の「異常マップ」の違い
- 「スコア比」の増減
- パラメータによる反応の変化

＊本教材はウェブカメラを用いた観察実験を目的としており、ベンチマークの完全再現ではありません。＊

NanoBanana2 画像異常検知 比較実験ツール

静的被写体の外観ベース異常検知（追加学習なし・教材用）

1. 正常登録 (ウェブカメラ) → 2. 連続比較検知 ⟶ 3. 異常マップ観察

PaDiM (統計モデル)

- 特徴抽出: ResNet-18 (ImageNet)
- 仕組み: パッチごとのガウス分布、マハラノビス距離

出力: 異常マップ、スコア比

PatchCore (メモリバンク)

- 特徴抽出: ResNet-18 (ImageNet)
- 仕組み: コアセット選択メモリ、最近傍距離

出力: 異常マップ、スコア比

DINOv2最近傍法 (基盤モデル特徴)

- 特徴抽出: DINOv2 ViT-S/14 (自己教師あり)
- 仕組み: 参照メモリ、コサイン距離

出力: 異常マップ、スコア比

*論文再現ではない教材実装 / AUROC等評価なし / ウェブカメラ入力 / 静止被写体対象 / 見た目の変化を検知

画像からの異常検知（PaDiM・PatchCore・DINOv2最近傍）

1. 実験目的＆異常の定義

正常登録した状態からの見た目の変化を、異常マップで観察する

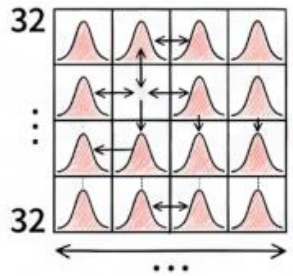
見た目の変化

正常登録した状態からの見た目の変化を、異常マップで観察する

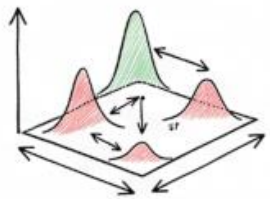
異常 = 異物・傷・汚れ・照明・ピント・影・位置ずれなど

2. 3方式の仕組み

A: PaDiM（統計モデル）

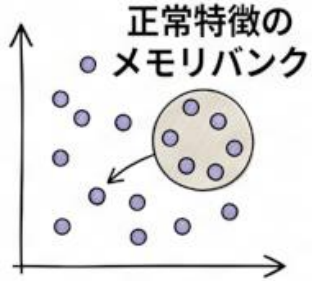


パッチ位置ごとにガウス分布

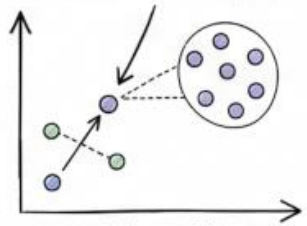


マハラノビス距離

B: PatchCore（メモリバンク）

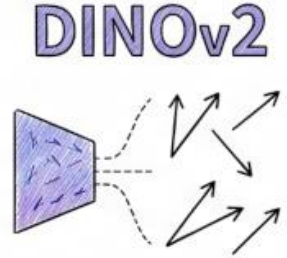


正常特徴のメモリバンク最近傍距離 (k)

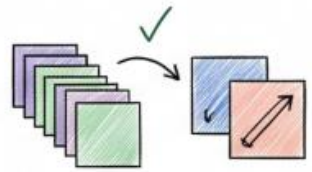


最近傍距離 (k)

C: DINOv2最近傍（基盤モデル特徴）



DINOv2特徴ファテクス特徴

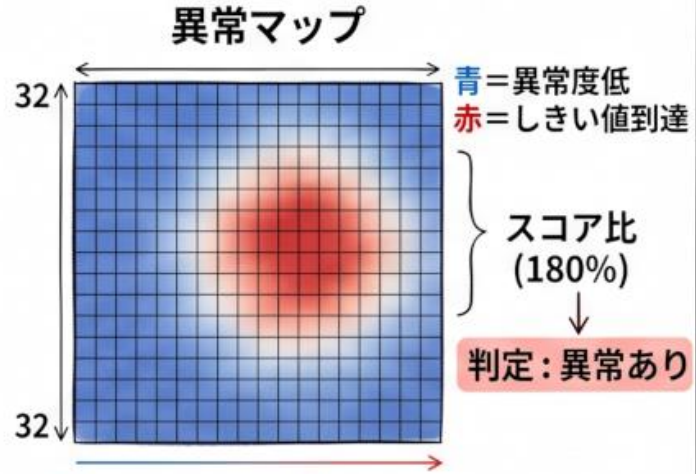


コサイン距離

共通: 追加学習なし。学習済み特徴抽出器と正常登録のみ。

3. 実験結果と考察

確認できたこと：異物・照明・位置ずれへの反応差、パラメータの影響



重要な注意 (プレゼンの要点)：

- スコア比：手法間の数値比較は不可 (相対値)
- 異常マップ: 32×32パッチ単位、赤は高異常度 (正確な境界ではない)
- 欠陥の理解ではなく「見た目の変化」への反応

4. 実験の限界

一般論としての優劣比較ではない
AUROC等の定量指標はない

5.低照度化と輝度補正が物体検出モデルの検出結果に与える影響の測定

1. 映像入力源

カメラデバイス　デバイス番号 0

動画ファイル　ファイルを選択

（未選択）

2. 照度条件（独立変数）

劣化ガンマ指数（中段の行）

0.60

劣化ガンマ指数（下段の行）

0.30

上段は基準条件として 1.00 に固定される

3. 検出器の設定（統制変数）

YOLO26n 確信度しきい値 conf

0.25

YOLO26n 推論入力の一辺 imgsz（画素）

640

参照集合との一致判定 IoUしきい値

0.50

4. 補正手法の設定（処置）

CLAHE クリップ制限値 clipLimit（第2列・第4列）

2.0

CLAHE タイル分割数（縦横共通、第2列・第4列）

8

AGCWD 重み付け指数 α（第3列・第4列）

0.50

照明マップ調整指数 γ_L（第5列）

0.80

ガイデッドフィルタ半径 r（画素、第5列）

15

5. 測定値の記録

☑ コンソールへ出力する

☐ CSVファイルへ記録する

記録先: measurement_log.csv

記録間隔（フレーム）

20

6. 実行制御

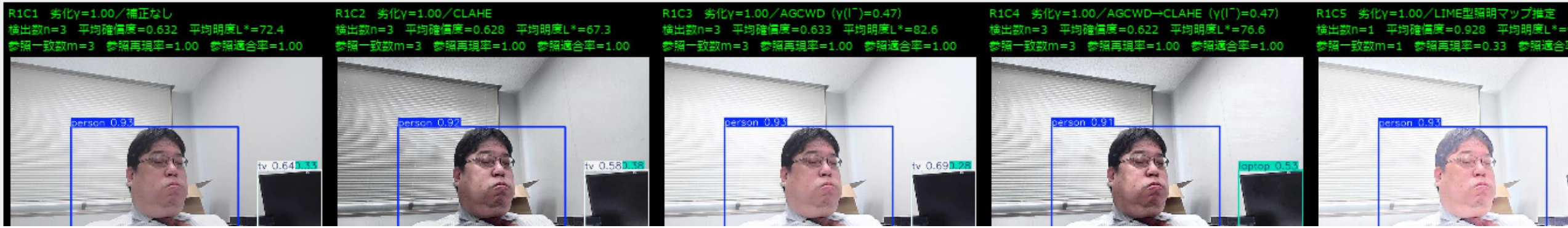
撮影開始　撮影停止

☐ 静止フレーム保持（同一フレームを反復処理する）

フレーム 242　1フレームの処理時間 229 ms（15条件のバッチ推論を含む）　参照集合の要素数 3

7. 測定結果（15条件を同時表示）

列（左から）：補正なし／CLAHE／AGCWD／AGCWD→CLAHE／LIME型照明マップ推定　　行（上から）：劣化ガンマ指数 1.00（基準条件）／中段／下段



実験探求の流れ：低照度化と輝度補正の影響測定

1. 入力と前提

入力：1枚の元画像
YOLO26nモデル

2. 実験条件の生成

輝度・コントラスト補正
（前処理）

		なし（対照）	ヒストグラム変換（CLAHE）	ガンマ変換（AGCWD）	AGCWD→CLAHE	照明マップ推定（LIME）
照度（ガンマ変換）	基準(1.00)	計15条件下で画像を生成				
	中					
	低					

計15条件

3. 測定・評価・出力

物体検出(YOLO26n)

15条件の画像を測定

測定指標
- 検出数 n, 平均確信度, 平均明度
- 参照一致数 m（基準R1C1と比較）, 参照再現率, 参照適合率

出力
- 15格子画像
- CSVログ

低照化と輝度補正の彩響の照維質る物体検出結果モデル結果を検定を測定

目的と構成

同一フレームに照度低下（ガンマ変換）と輝度補正を組み合わせ、物体検出結果の変化を測定。

構造:
5列×3行の格子。検出映像、検出画像やと測定をある。列の内中にで同じ劣化画像をある、共有

使用技術とAIモデル

- 照度劣化: ガンマ変換
- 輝度補正手法
  1. 補正なし
  2. CLAHE (Contrast Limited Adaptive Histogram Equalization)
  3. AGCWD (Adaptive Gamma Correction with Weighting Distribution)
  4. AGCWD→CLAHE
  5. LIME型照明マップ推定
- 物体検出モデル: YOLO26n (Ultralytics, 2026)
  - 特徴
  - NMSを用いない一段物体検出モデル
  - 最大300個の検出を直接出力
  - COCO mAP50-95: 40.9
  - パラメータ数: 2.4M
- 学習済みデータ
  - COCO 2017（80クラス）

測定結果の構成と項目

行：照度条件
(劣化ガンマ 1.00 / 中段 / 下段)

列：補正条件
(補正なし / CLAHE / AGCWD / AGCWD→CLAHE / LIME型)

測定項目

- 検出数 n
- 平均確信度
- 参照一致数 m
- 参照再現率
- 参照適合率
- 平均明度 L*

実験コンセプト

暗さと補正が物体検出に与える影響を測る

同一フレームから15条件を生成

すべての条件を常に同時に表示

1フレーム → 3照度 → 5補正

▼

15条件

▼

YOLO26n 一括推論

使用した手法

物体検出モデル

YOLO26n

(COCO学習済み、NMSなし)

照度変化 (行方向)

劣化ガンマ変換

(1.00 / 中 / 下)

輝度補正 (列方向)

補正なし (対照条件)

CLAHE

(局所コントラスト強調)

AGCWD

(統計依存トーン変換)

LIME型

(照明マップ推定)

測定値と指標

検出数 n: 確信度しきい値以上の個数

基準一致数 m: 左上のマスを参照集合とする

参照再現率: 基準条件の出力をどれだけ保っているか

平均明度 L*: 入力画像のCIELAB明度