

ジェスチャープレゼンター

Riki ISHIBASHI

Ren UETA

Teppei OKANO

Daichi SEGAWA

Shoya WATANABE

はじめに

上手なプレゼンテーションとは？

- ✓ ジェスチャーや表情の変化を取り入れている
- ✓ 立ち位置や姿勢が整っている など

同じ資料でも話し方を工夫すればより良いものに

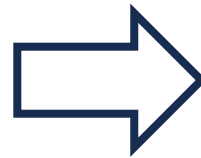
しかし..

学生の発表のほとんどは単調になる傾向
プレゼンにジェスチャーを挟む恥ずかしさ
上手なプレゼンを行うことは困難

目的

発表者と聴講者の両者が楽しめるプレゼン支援ツールの開発

体の動きでスライドを操作
静的から動的な発表へ



興味・関心・説得力の増加
プレゼンテーションの円滑化
プレゼン時の緊張感の緩和

開発内容

□ 発表者のジェスチャーを認識しプレゼン内容が動作

- 指の動きで画面を切り替え
- 発表者の動作で拍手などエフェクトの表示
- 表情を認識し、それに応じたエフェクトの表示 など

可能であれば実装したいオプション 例

- 発表時間や声量に応じたアクション
声量が小さい場合に発表者への注意を促す
- 聴講者の表情やリアクションを随時フィードバック
観衆の反応をデータでうかがえる

技術要素

□ ジェスチャー・表情の認識

- PC内蔵カメラ・USBカメラ

□ 音声の入力

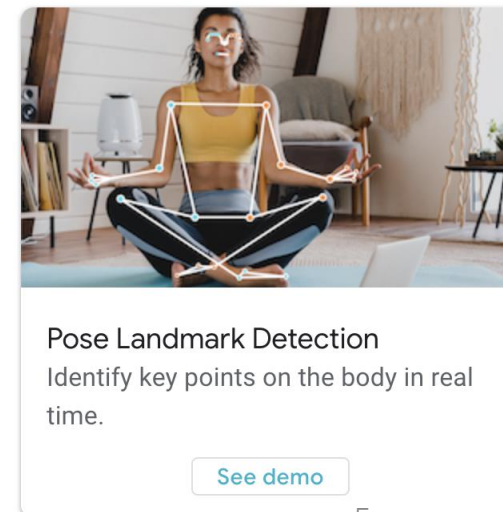
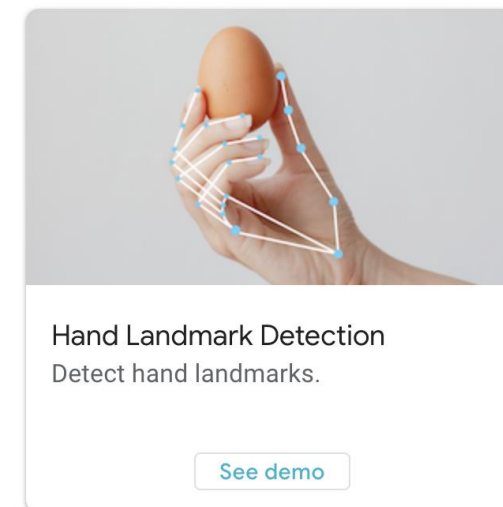
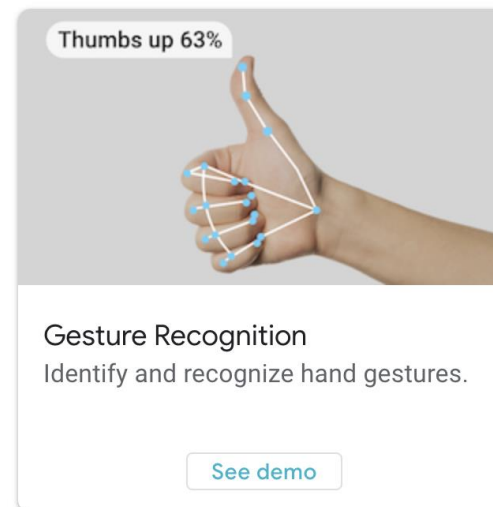
- マイク

□ 効果音等の音出し

- スピーカー

□ 人体認識：Google MediaPipe

- 機械学習ライブラリ
- 全身・ジェスチャー・手の形などを認識可能

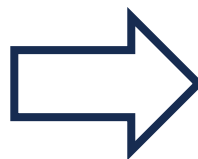


技術要素

□ Google MediaPipe

入力画像に対し、機械学習モデルを用いてタスクを処理

- オブジェクト認識
- ジェスチャー認識
- 画像処理
- 顔検出
- 全身関節座標の検出



今回は

- 全身関節座標の検出
- ジェスチャー認識

を使用

技術要素

□ YOLO

- 機械学習ライブラリ
- 表情判定のモデルを作成、実装
- 発表時の表情フィードバックに使用

□ Swiper.js

- スライドの制御に使用

□ Chart.js

- プレゼンテーションのフィードバックをグラフ描画によって行う

□ Next.js Typescript

- アプリのフロントエンド用フレームワーク
- Mediapipe・Swiperを制御

□ Go(Gin)

- APIサーバー
- 認証

□ MySQL

- ユーザ
- プレゼンテーションスライド

YOLOで用いたデータセット

□ データセット：表情画像データ AFFECTNET

- 感情値、覚醒の強さについて手動でラベル付けされたデータ

- ラベル： neutral, happy, angry, sad, fear, surprise, disgust, contempt
(無表情 幸せ 怒り 悲しみ 恐怖 驚き 嫌悪 軽蔑)

- Train：約17000枚

- Val：約5500枚

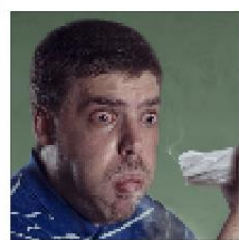
- 解像度：96×96

- メモリ削減のため低解像度解像度

YOLOで用いたデータセット



ffhq_147.png



ffhq_148.png



ffhq_149.png



ffhq_150.png



ffhq_162.png



ffhq_163.png



ffhq_164.png



ffhq_165.png



ffhq_158.png



ffhq_159.png



ffhq_160.png



ffhq_161.png



ffhq_174.png



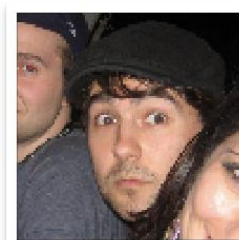
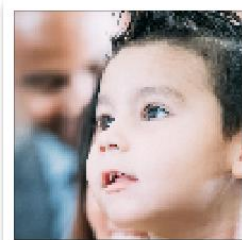
ffhq_175.png



ffhq_176.png



ffhq_177.png



YOLO 学習

- モデル：v8small(モデルサイズ要検討)
- エポック数：200
- 解像度：640×640
 - (96×96) では入力画像サイズが異なり検出不可

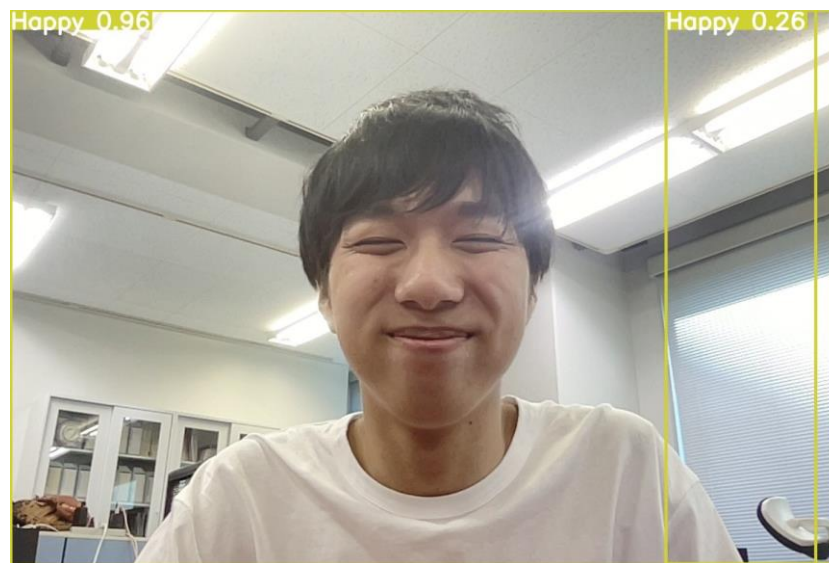
allクラスの学習結果 (val)

P	R	mAP50	mAP50-95
0.736	0.767	0.83	0.83

YOLO 検出結果画像



Neutral
73%



Happy
96%



Disgust
91%

システムの特徴

□ 類似製品はなし

□ 類似研究・他研究との比較

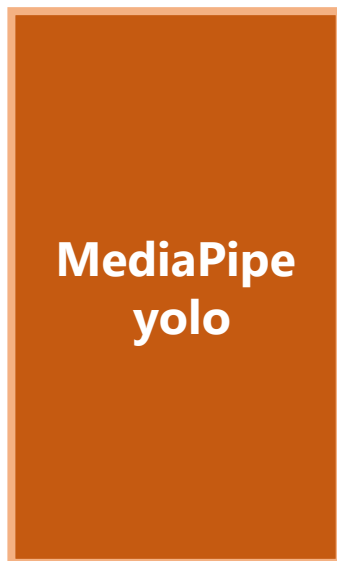
- Kinectやセンサなどの特殊デバイスを必要としない
- エンタメ要素を付加する独自のソリューション

□ USBカメラやスピーカーなどのIoTデバイスを活用

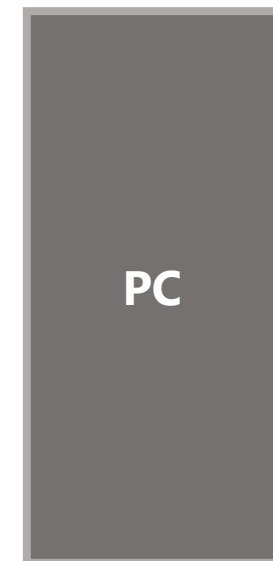
□ 機械推定による姿勢推定を実施

システム全体像

発表者の映像を入力



発表者の音声を入力



システム内のAIの活用箇所

□ AI

■ Google MediaPipe：ジェスチャー・姿勢検出

- スライド操作

■ YoLo：表情認識

- 発表時の表情フィードバック

スケジュール

	5月					6月				7月			
各水曜日	3	8	15	22	29	5	12	19	26	3	10	17	24
アイディア練って、発表	アイディア練り												
姿勢推定						姿勢推定							
スライド切り替え機能の作成										スライド切り替え			
モーションに合わせた効果音機能										オーディオ機能			
モーションに合わせた画面効果											エフェクト作成		
テスト											テスト		
発表準備・成果発表												発表準備	

イマココ







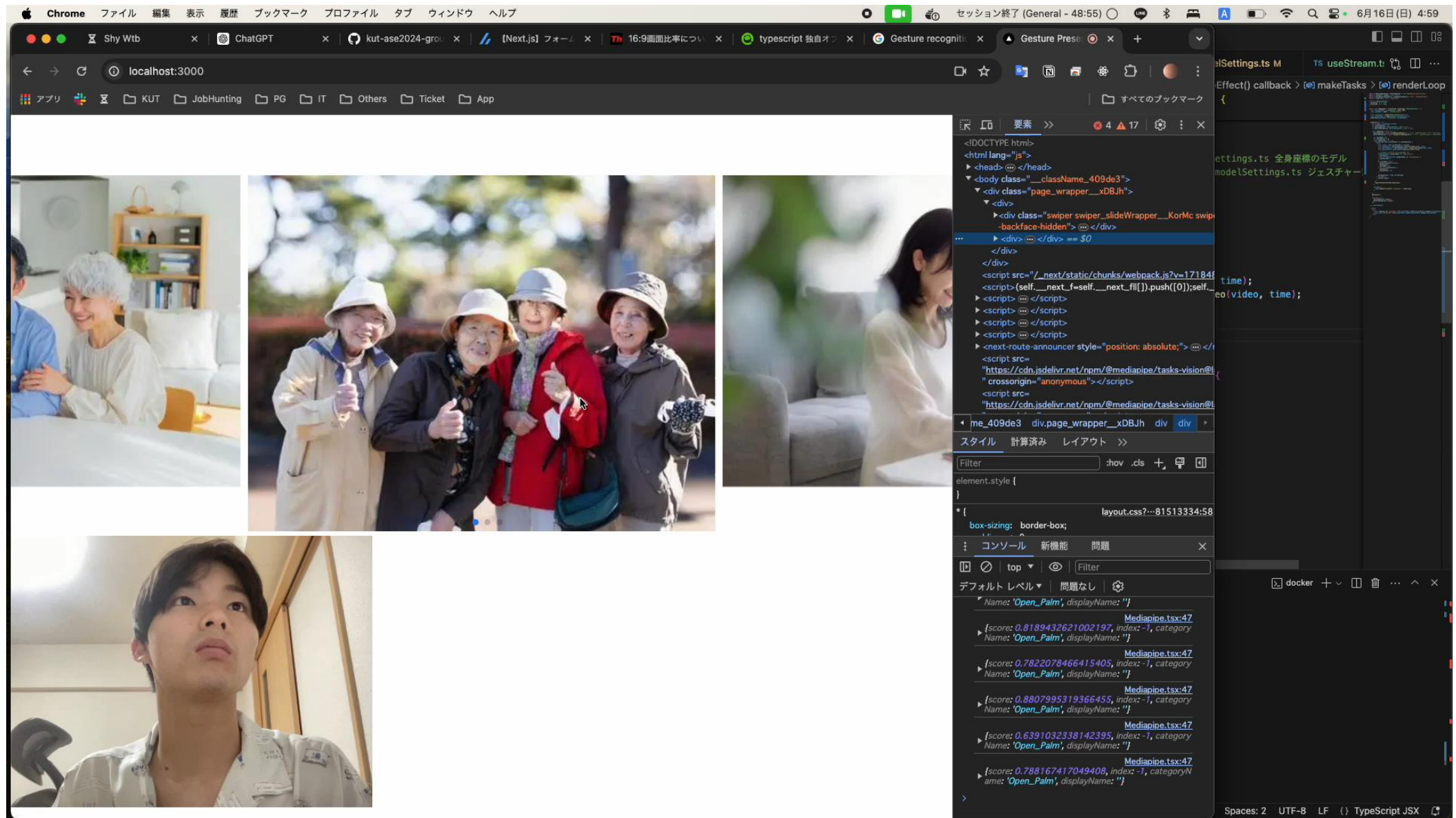
スライド切り替え

各自の分担

- 植田：インフラ構築・バックエンド開発
- 石橋：表情認識モデル作成
- 渡部：フロントエンド開発
- 岡野：フロントエンド開発
- 瀬川：フロントエンド開発・関連資料作成

現在の進行状況

- Docker, Git による環境構築 
- Mediapipeによるジェスチャー認識 
- ジェスチャーによるスライド前後操作 
- 表情認識モデルの作成 



手のひらをパーの形：ページ送り

人差し指を立てる：ページ戻し



右手を挙げる：ページ送り

左手を挙げる：ページ戻し

検討事項

- プレゼンの評価・フィードバック
- ジェスチャー認識の精度とスライド操作の利便性
- 効果音や、エフェクトの追加
- セキュリティの向上

今後の見通し

- 表情認識を用いたフィードバックの機能の実装
- ジェスチャー認識の適切な閾値の設定
- スライド動作のレパートリーの追加
- ログイン機能の実装
- 頑張って開発する

前回のフィードバックへの返答

□ 内蔵カメラと外部設置カメラの写りの差への対応について

- 内蔵カメラの使用を前提としており、
外部カメラの使用は現在考えておりません。

□ 緊張によって真顔になっている場合に、 笑顔になってなどのフィードバックを行えば嬉しいかも

- CNNでの表情フィードバックの実装に取り組んでいます。

まとめ

- 発表者・聴講者が退屈しないプレゼン支援ツールを開発
- ジェスチャーなどによるスライド操作で動的な発表へ
- エンタメ要素を付加し、華やかなプレゼンに
- 以降は締め切りを視野に入れながらオプションへの取り組みも