

# 知能情報論 最終課題

37-246392 豊田 圭将

2024 年 8 月 2 日

## 1 はじめに

本レポートでは、YouTube の動画分析ツールの実装について説明する。このツールは、動画の情報取得、自動生成された字幕抽出、OCR 処理、および ROUGE-L スコアの計算と可視化を行うものである。

## 2 実現した知的機械の概要と考えられる応用先

本プロジェクトで実現した知的機械は、YouTube の動画コンテンツを自動的に分析し、テキストデータの一致度を評価するシステムである。主な機能は動画のメタデータ収集、自動生成された字幕の抽出、動画フレームからの OCR 処理、自動生成された字幕と OCR 結果の比較分析、および分析結果の可視化である。

応用先としては、動画の字幕品質評価、コンテンツモデレーション、悪意のあるテキスト付き動画の検出などが挙げられる。

## 3 利用した手法の説明

本システムでは、以下の主要な手法を利用している：

- ・ YouTube Data API：動画のメタデータ取得
- ・ YouTube Transcript API：自動生成された字幕の取得
- ・ Google Cloud Vision API：OCR 処理によるテキスト抽出
- ・ ROUGE-L：自動生成された字幕と OCR 結果のテキスト一致度評価
- ・ 並列処理：ThreadPoolExecutor による処理速度向上

## 4 工夫した点の説明

本システムの実装において、以下の点で独自の工夫を行った：

- ・ 動的な OCR 処理範囲の決定：動画種類に応じて OCR 処理領域を調整
- ・ 適応的な OCR 処理間隔：字幕のタイムスタンプに応じてフレームを抽出
- ・ マルチスレッド処理の最適化：CPU コア数に応じて最適なスレッド数を決定
- ・ エラー復旧メカニズム：指数バックオフアルゴリズムによる再試行
- ・ チャンネル別分析機能：チャンネルごとの ROUGE-L スコア集計と比較分析

## 5 利用した実データの説明

本研究では、東京都知事選挙に関する YouTube 動画 500 本（通常動画 342 本、ショート動画 158 本）を分析対象とした。通常動画では最初の 5 分間、ショート動画では 1 分間に対して処理を行った。

## 6 得られた結果

システムの実行結果から、以下の定量的評価と知見が得られた：

### 6.1 ROUGE-L スコアの分布

- ・ 全動画の平均 ROUGE-L スコア：0.37
- ・ 通常動画の平均 ROUGE-L スコア：0.36
- ・ ショート動画の平均 ROUGE-L スコア：0.41

予想に反し、ショート動画の方が通常動画よりも ROUGE-L スコアが高い傾向が見られた。

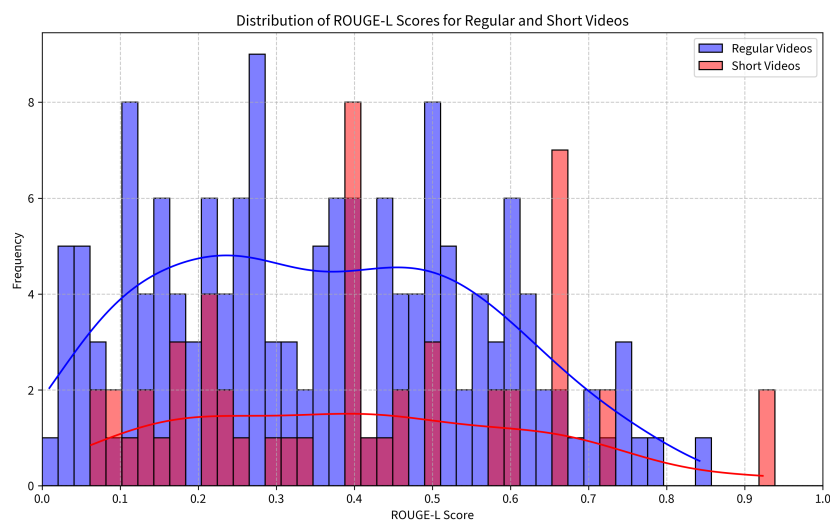


図 1 通常動画とショート動画の ROUGE-L スコア分布の比較

### 6.2 再生回数と ROUGE-L スコアの相関

再生回数と ROUGE-L スコアの間にはほとんど相関が見られなかった（Spearman’ s rank correlation：0.049）。

### 6.3 動画の長さと ROUGE-L スコアの関係

動画の長さ（秒）と ROUGE-L スコアの間には弱い負の相関が見られた（Spearman’ s rank correlation：-0.19）。

### 6.4 チャンネル別の分析

チャンネル間で字幕の品質に差があることが明らかになった。

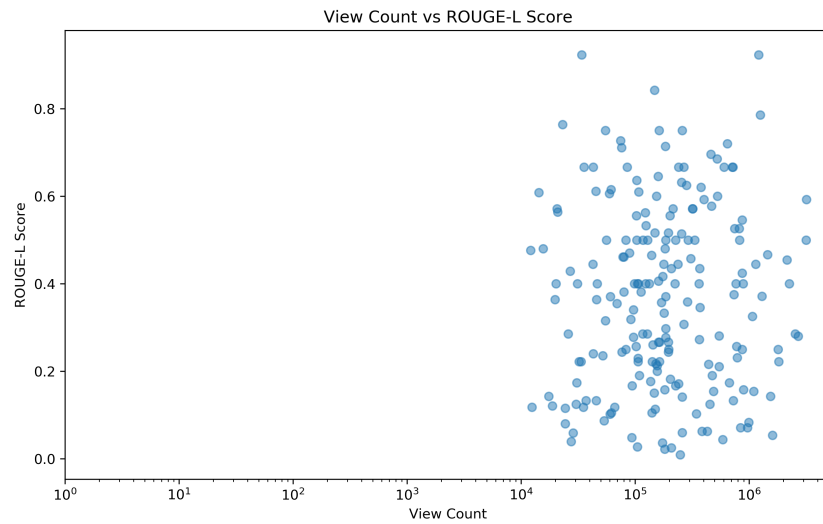


図 2 再生回数と ROUGE-L スコアの関係

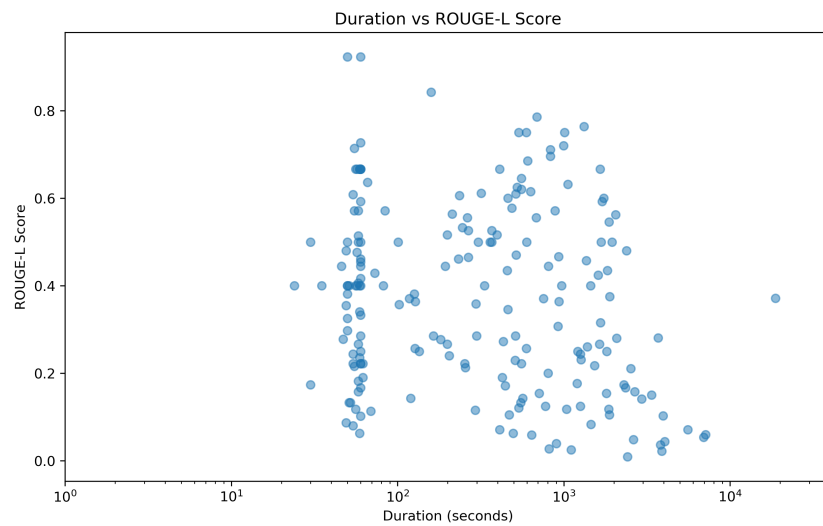


図 3 動画の長さと ROUGE-L スコアの関係

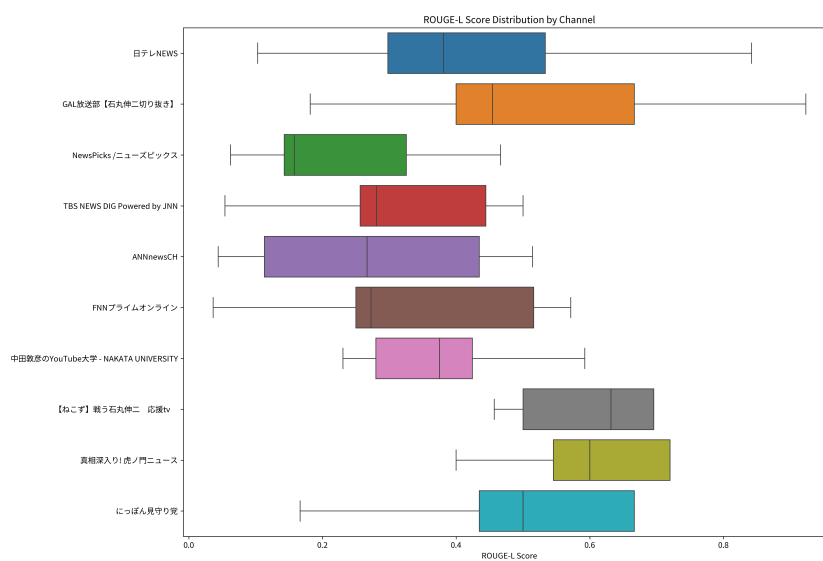


図 4 主要 10 チャンネルの ROUGE-L スコア分布

## 7 結論

本プロジェクトでは、YouTube の動画分析ツールを実装し、500 本の実際の動画データを用いて評価を行った。得られた主な結果は以下の通りである：

- ・ ショート動画の方が通常動画よりも ROUGE-L スコアが高い傾向が見られた。これは、ショート動画の方が手動で字幕を付ける手間が少なく、付与されにくいことを示唆している可能性がある。
- ・ 動画の長さと言語品質には弱い負の相関が見られた。これは、動画が長くなるほど字幕の品質が若干低下する傾向があることを示唆している。
- ・ 再生回数と言語品質にはほとんど相関が見られなかった。これは、動画の人気度と言語品質には直接的な関連がない可能性を示している。
- ・ チャンネルによって字幕の品質に大きな差があることが明らかになった。これは、チャンネルごとの字幕生成や管理方針の違いを反映している可能性がある。

今後の課題としては、より多様なジャンルの動画での評価、言語別の分析、時系列での字幕品質の変化の追跡などが挙げられる。本研究の成果は、動画コンテンツの品質向上、自動字幕生成システムの改善、効果的な動画マーケティング戦略の立案など、様々な分野での応用が期待できる。

## 8 使用したライブラリとコードの実行方法

本プロジェクトで使用した主要なライブラリは以下の通りです：

- ・ numpy
- ・ pandas
- ・ matplotlib
- ・ japanize-matplotlib
- ・ seaborn
- ・ opencv-python
- ・ rouge-score
- ・ yt-dlp
- ・ youtube-transcript-api
- ・ google-oauth
- ・ google-api-python-client
- ・ google-cloud-vision

コードの実行方法や詳細な実装については、以下の GitHub リポジトリを参照してください：

[https://github.com/KeisukeTTTT/transcript\\_analysis.git](https://github.com/KeisukeTTTT/transcript_analysis.git)

このリポジトリには、本プロジェクトで利用したすべてのソースコード、データ処理スクリプト、および結果の可視化スクリプトが含まれています。README ファイルに従って環境をセットアップし、スクリプトを実行することで、本レポートで示した分析を再現することができます。