

# ディープフェイク検知システム SYNTHETIQ VISION -

安田裕介 \*<sup>1</sup>、古橋良彦 \*<sup>2</sup>、山岸順一 \*<sup>1, 3</sup>、越前功 \*<sup>2, 3</sup> (\*<sup>1</sup> コンテンツ科学研究系、\*<sup>2</sup> 情報社会相関研究系、\*<sup>3</sup> シンセティックメディア国際研究センター)

## どんな研究？

- SYNTHETIQ VISION は、生成 AI によって生成された画像か本物の画像かを見分けるシステムです。
- 特に生成 AI の悪用の 1 つであるディープフェイクの検知ができます。
- ディープフェイク検知を Web API で操作できるようになっています。

## 何がわかる？

- 生成 AI を悪用した偽情報への対策ができます。
- 人間には本物と見分けがつかない生成画像であっても、検知システムは見分けることができます。

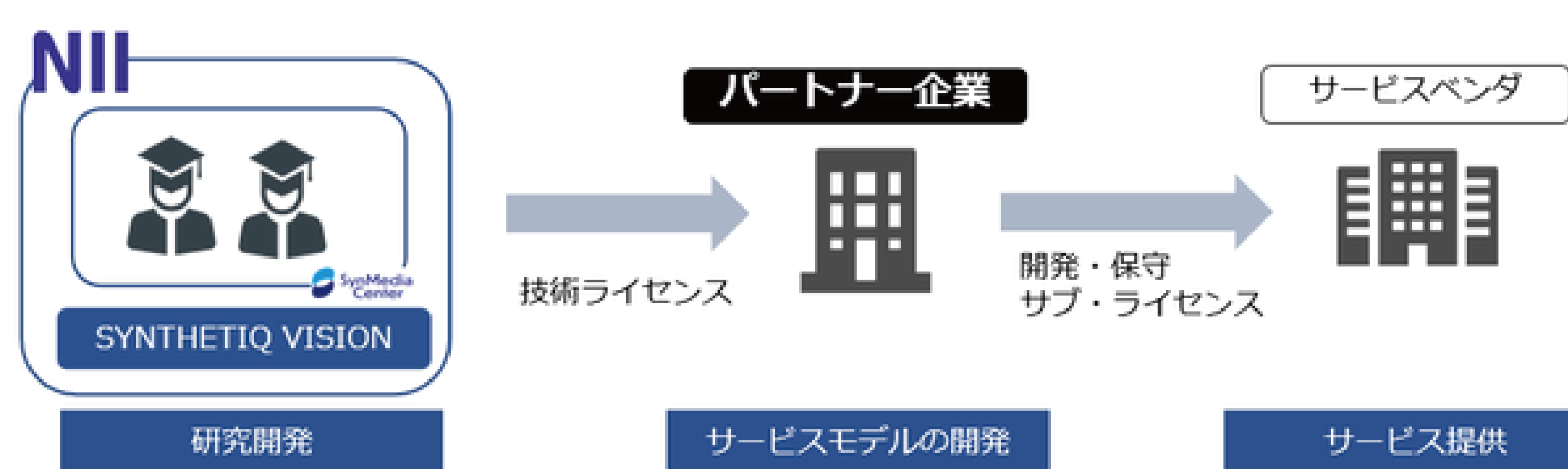
## 背景・目的

### 背景

- 生成 AI の性能は劇的に向上しています。
- 生成 AI を悪用したディープフェイクの被害は実際に出始めており（例：詐欺）、将来はより脅威が高まることが予想されます。
- 私たちはディープフェイク検知技術の研究に長年取り組んできました。
- 私たちのディープフェイク検知技術を、より広く利用できるようにシステム化したものをデモします。

### 目的

- 生成された動画画像かどうかを検知することにより、偽情報に対する対策を可能にする。
- ディープフェイク検知技術を広く使いやすい形式で提供する。



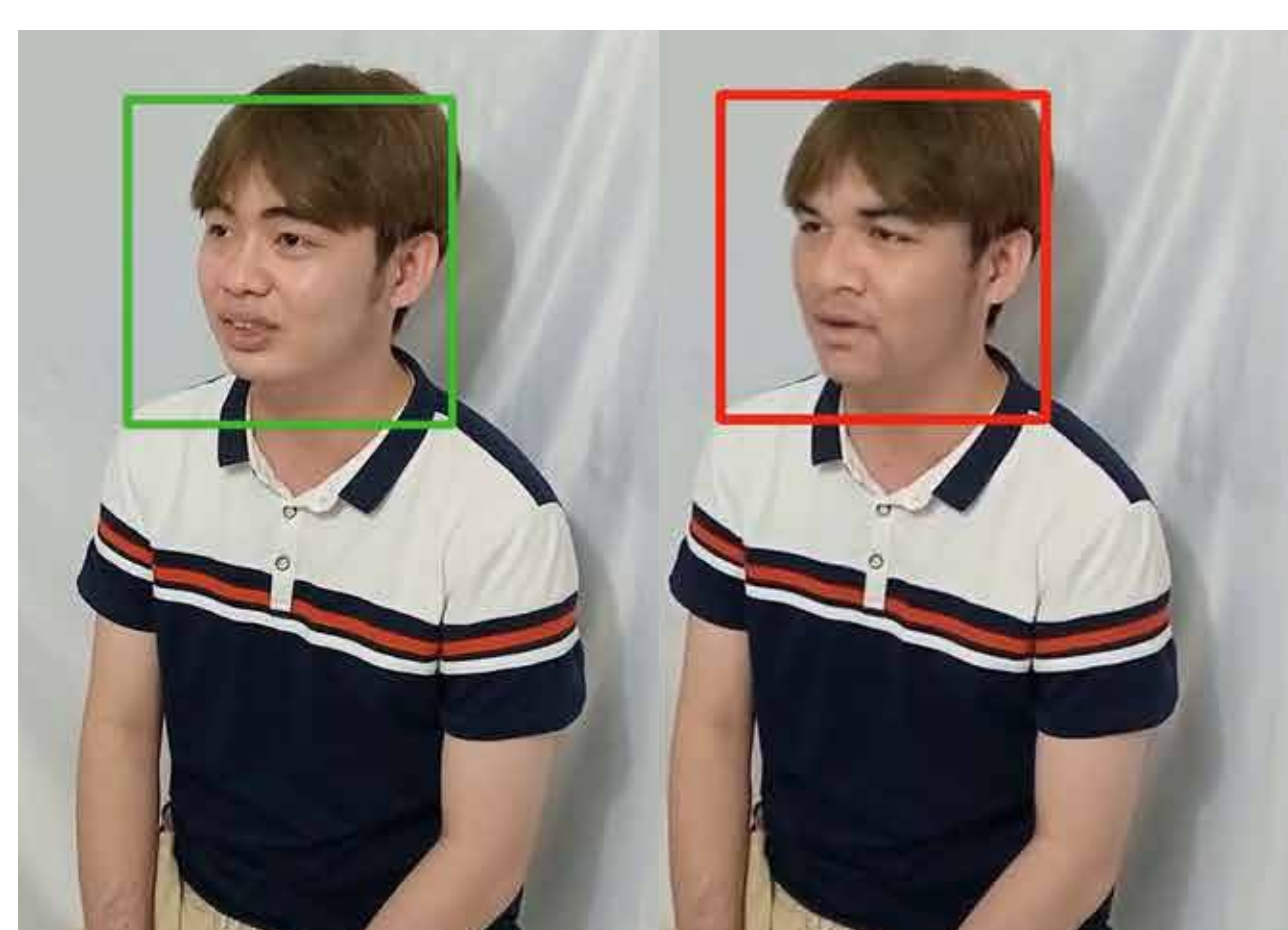
## 研究内容（方法・結果・結論）

### 何ができるのか

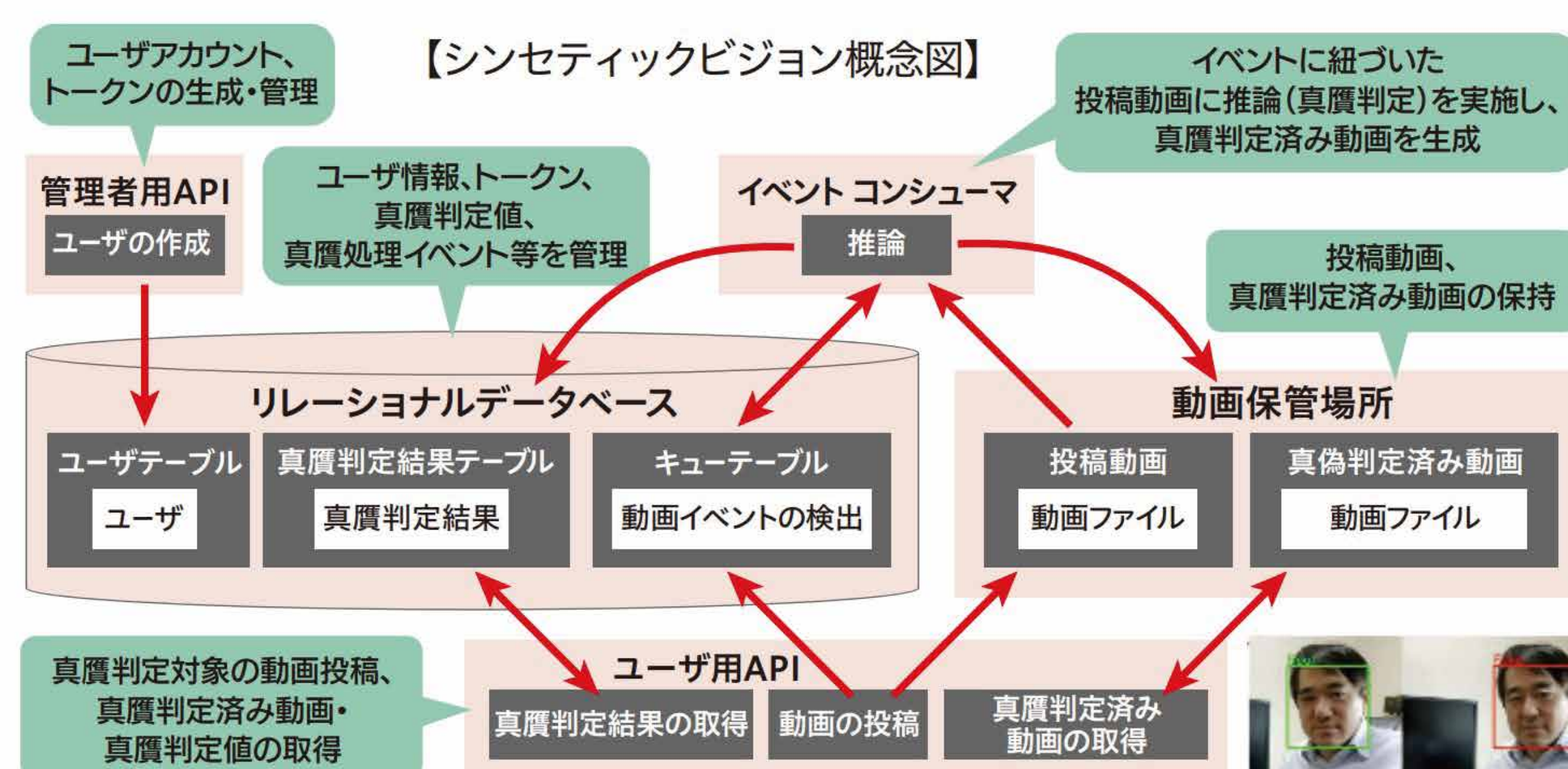
- Q: どうやって見分けているのですか？  
A: 深層学習を使用し、大量の偽画像とリアル画像の判別を学習しています。
- Q: 何を元に判定しているのですか？  
A: 偽画像特有のアーティファクトを元にモデルが判定しています。
- Q: 精度はどのくらいなのですか？  
A: 私たちの評価データにおいて、90% 以上の精度があります。
- Q: 学習した人間しか判定できないのですか？  
A: 学習していない人間でも判定できます。
- Q: 顔画像以外（自然画像等）は判定できないのですか？  
A: 顔画像のみです。自然画像は将来対応する可能性があります。

### どのように実現しているのか

- API サーバーとキュー消費プロセスの 2 つのサブシステムでシステムを構成。
- API サーバーで検知リクエストを受け、検知リクエストイベントをキュー消費プロセスで非同期に処理。



図：判定結果の可視化



図：SYNTHETIQ VISION の仕組み

### まとめ

- 生成 AI によって生成された画像か本物の画像かを見分けるシステムを開発しました。
- 人物の顔を含んだ画像や動画のディープフェイク検知に有用です。
- Web API で利用できるため、利用しやすい形で提供しています。



詳細はこちら