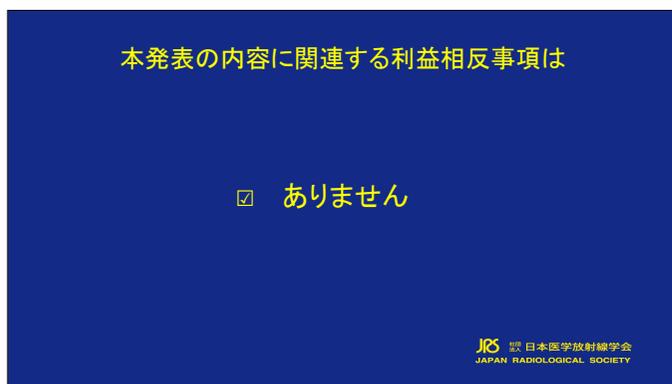
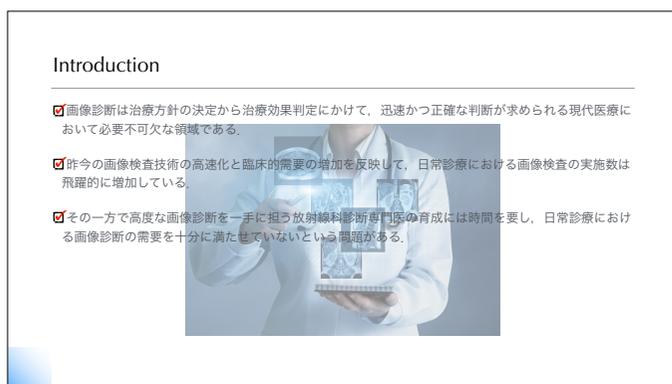




皆さんこんにちは。信州大学医学部画像医学教室の山田と申します。  
ここでは、我々の研究室と、日本ケイデンス社との共同研究として進めている、「視線追跡型VRヘッドマウントディスプレイと音声認識コマンド操作によるハンズフリーVR画像診断環境構築ならびに人工知能学習用教師ラベル作成の試み」についてご紹介させていただきます。



利益相反はありません。



▲画像診断は治療方針の決定から治療効果判定にかけて、迅速かつ正確な判断が求められる現代医療において必要不可欠な領域となっています。  
▲昨今の画像検査技術の高速化と臨床的需要の増加を反映して、日常診療における画像検査の実施数は飛躍的に増加しています。  
▲その一方で高度な画像診断を一手に担う、放射線科診断専門医の育成には時間を要し、日常診療における画像診断の需要を十分に満たせていないという問題があります。

## Introduction

☑この問題を解決するために、放射線科診断専門医の画像認識力を高め診断精度・効率を向上させる仮想現実 (Virtual Reality; VR) 技術の応用や、放射線科診断専門医の診断支援を行う人工知能 (Artificial Intelligence; AI) の開発に注目が集まっています。

☑VR環境下では外部環境の視認が困難となるため、直感的かつ低遅延な非用手的デバイス操作が重要となってくる。

☑一方、効率の良い人工知能学習用教師ラベル作成のためにも、直感的なインターフェースの構築は必要不可欠である。



▲この問題を解決するために、放射線科診断専門医の画像認識力を高め診断精度・効率を向上させる仮想現実 (Virtual Reality; VR) 技術の応用や、放射線科診断専門医の診断支援を行う人工知能 (Artificial Intelligence; AI) の開発に注目が集まっています。

▲しかしながら、VR環境下では外部環境の視認が困難となるため、直感的かつ低遅延な非用手的デバイス操作が重要となってきます。

▲一方、効率の良い人工知能学習用教師ラベル作成のためにも、直感的なインターフェースの構築は必要不可欠な要素です。

## Introduction

☑今回我々は、直感的かつ効率的な人工知能学習用教師ラベル作成を行うデバイスとして視線追跡型VRヘッドマウントディスプレイに注目し、直感的かつ遅延のない非用手的なデバイス操作法として音声認識コマンド操作ボードを用いたハンズフリーVR画像診断環境の構築に成功した。

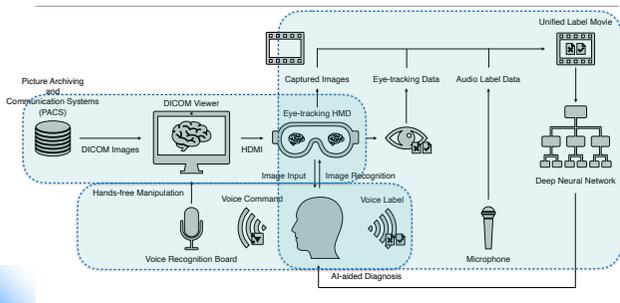
☑本発表では、新規開発したハンズフリーVR画像診断環境を医学生の画像解剖教育に役立てながら人工知能学習ラベル作成に応用した具体的事例について紹介する。



▲今回我々は、直感的かつ効率的な人工知能学習用教師ラベル作成を行うデバイスとして視線追跡型VRヘッドマウントディスプレイに注目し、直感的かつ遅延のない非用手的なデバイス操作法として音声認識コマンド操作ボードを用いたハンズフリーVR画像診断環境の構築に成功しました。

▲本発表では、新規開発したハンズフリーVR画像診断環境を医学生の画像解剖教育に役立てながら人工知能学習ラベル作成に応用した具体的事例について紹介いたします。

## Materials & Methods

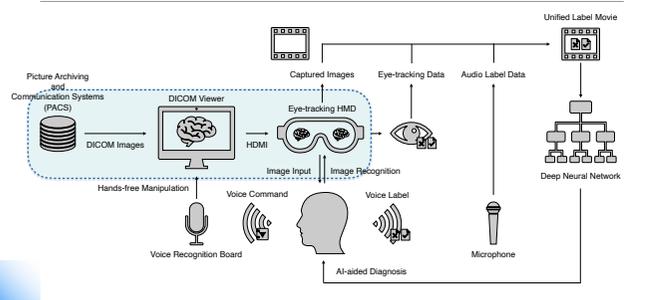


今回開発したハンズフリーVR読影環境の概要を示します。

本環境は、▲画像提示、▲ラベル作成、▲音声認識操作という主な3つの構成要素から成り立っています。

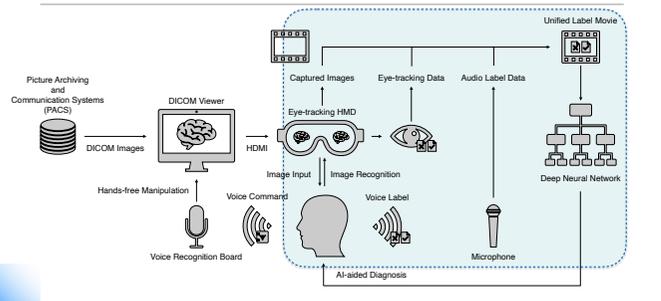
▲それぞれの構成要素について、順に説明していきます。

### Materials & Methods: Image Presentation



画像提示部では、DICOMビューアーから病院内のPACSシステムにアクセスし、匿名化したDICOMビューアーのモニター出力を視線追跡型ヘッドマウントディスプレイにミラーリングして観察者に画像提示を行います。PACSシステム外への医用画像データの持ち出しが生じないため、個人情報保護の観点からも優れています。

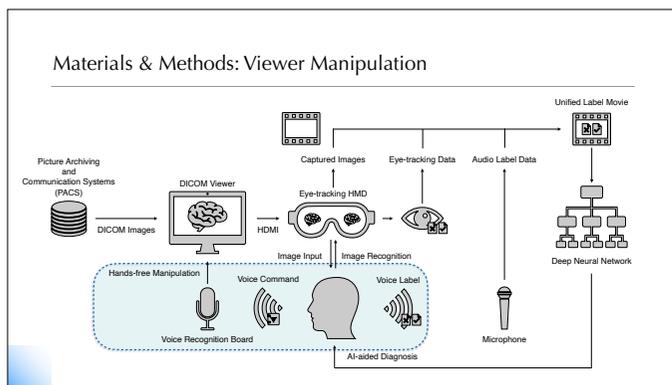
### Materials & Methods: Label Creation



ラベル作成部では、DICOMビューアーからのキャプチャー画像と、視線追跡型HMDによる視線情報、そしてラベル音声を統合したラベル動画を生成します。画像キャプチャーは毎秒60フレーム、視線追跡は毎秒120フレームで行うため、理論的には毎秒60組の人工知能学習用の教師ラベルを生成することが可能となります。

視線には生理的な揺らぎが存在することから、同一キャプチャ画像に対して自然な教師ラベルの augmentation を実現することが可能です。

また、視線や音声を活用することによって、従来のマウスやキーボードなどのユーザーインターフェースと比べて、より直感的に教師情報を入力できることが利点です。



視線追跡型HMDは非常に高精度の視線追跡情報を得ることができますが、マウスやキーボードなどの従来のユーザーインターフェースではHMD装着下のDICOMビューアーの操作は困難であるため、音声認識コマンド操作を導入します。一方、コマンド操作のみならずラベル情報も音声で入力するため、ここではラベル音声とコマンド音声を識別可能な高精度の音声認識技術が必要となります。

それでは、このハンズフリーVR読影環境を用いた臨床応用の一例を動画でご紹介します。

ここでは、医学部3年生の学生を対象に、CT画像を用いた画像解剖実習を行いながら、人工知能学習用の教師ラベルをあわせて作成しています。







音声認識の精度は高く、  
未登録音声コマンドや雑音による誤動作はほとんど生じない



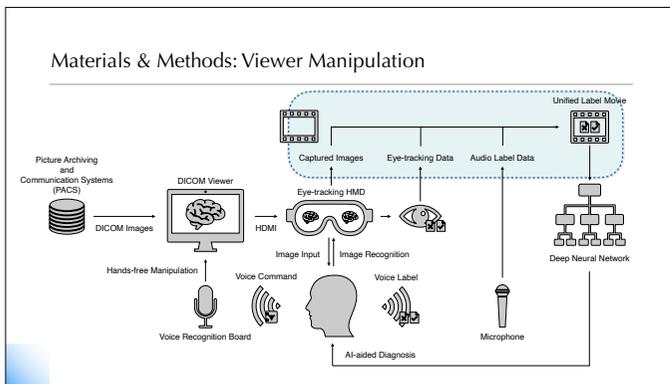
音声入力によるラベル指定と、視線入力による位置指定の  
組み合わせにより、効率的なラベル作成が可能となる



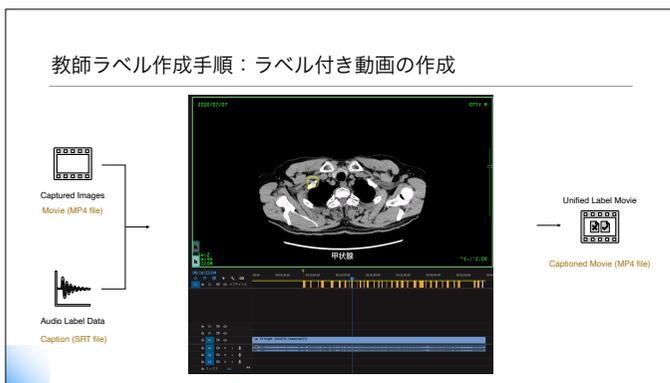
音声認識コマンド操作は不特定語言に対応しており、  
学生と教官がそれぞれ自由に操作することができる



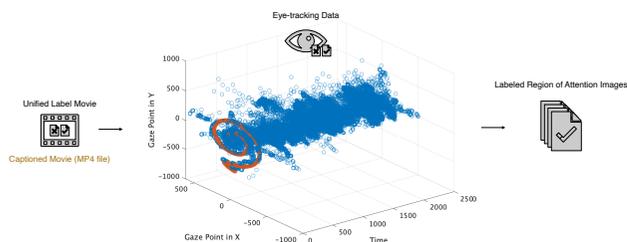
次に、このように得られた各種データを統合して人工知能学習用ラベルを作成する手順をご紹介します。



まずはじめに、ヘッドマウントディスプレイに表示された画像のキャプチャー動画と、教師ラベルを指定する音声データをキャプション情報として統合し、標的とする教師ラベルが動画中に存在する時間を特定します。



### 教師ラベル作成手順：ラベル付き注視領域 (ROA) 画像の作成



次に、キャプション付き動画と視線追跡データを統合し、動画中に教師ラベルが存在する部位を特定します。  
 このようにして、キャプチャー動画から標的とする教師ラベルを時間的・空間的に特定し、最終的な教師ラベル付きの ROA 画像 (Region of Attention) 画像を作成し、人工知能学習用データに活用します。

### 作成された実際のラベル付き注視領域 (ROA) 画像



実際に作成された ROA 画像を供覧します。今回の実習では、平均して約 40 分の作業で 50 解剖部位、1 解剖部位あたり約 48 秒でラベル作成が可能でした。

### 実習前後の医学生画像解剖理解の変化

教官の指定した50の解剖学的部位を全身非造影CT画像 (2.5 mm 厚、) から選び出す課題を実習前後に実施

正答率	学生A	学生B	学生C	平均
実習前	18% (9/50)	26% (13/50)	40% (20/50)	28% (14/50)
実習後	74% (37/50)	88% (44/50)	86% (43/50)	82.3% (41.3/50)

P = 0.0072 (検定)

さらに、実習前後における医学生の画像解剖理解の変化をみてみますと、実習前と比較して、実習終了時には有意に画像解剖理解が向上したことがわかります。

## Perspective

☑視線追跡型ヘッドマウントディスプレイと音声認識コマンド操作ボードを組み合わせたハンズフリー画像診断環境を用いて、医学生の画像解剖教育を行いながら効率的な人工知能学習ラベルの作成が可能であった。

☑今後、実際の放射線科診断専門医の読影業務への応用、人工知能画像診断支援システムの実用化に取り組んでいきたい。



いかがでしたでしょうか？

ここでは、視線追跡型ヘッドマウントディスプレイと音声認識コマンド操作ボードを組み合わせたハンズフリーVR画像診断環境を用いて、医学生の画像解剖教育を行いながら効率的な人工知能学習ラベルを作成する試みをご紹介しました。

今後、本環境を利用した実際の放射線科診断専門医の読影業務への応用、人工知能画像診断支援システムの実用化に取り組んでいきたいと考えています。

## 使用機材

☑視線追跡型HMD：FOVE0 (株式会社FOVE, 東京)  
▶視線追跡データ解析ソフトウェア  
▶自作 Unity ソフトウェア (Unity Technologies, USA)  
▶自作 MATLAB ソフトウェア (Mathworks, USA)

☑音声認識コマンド操作ボード：共同研究用試作機 (日本ケイデンス・デザイン・システム社, 横浜市)  
▶音声認識エンジン：Truly Handsfree (センサリージャパン株式会社, 横浜市)

☑DICOMビューア：EV Client (PSP 株式会社, 東京)

☑HDMIキャプチャー：4K60 PRO (Elgato, USA)  
▶キャプション動画作成ソフトウェア：Premiere Pro 2021 (Adobe Inc., USA)

本日よりご紹介した研究発表において使用した機材はご覧の通りです。

本発表の内容についてご質問等がございましたら、信州大学医学部画像医学教室、山田までご連絡ください。

閲覧いただき、誠にありがとうございました。