

Multimodal Data Integration in Medicine

Yuhan Wang

Center for Robotics Research, Department of Engineering,
King's College London

The Alan Turing Institute

Email: yuhan.4.wang@kcl.ac.uk

Table of Contents

- Introduction on Multimodal-Data Integration
- Alzheimer's Disease Detection

Multimodal Data

Google DeepMind Google Research

2024-04-29

Capabilities of Gemini Models in Medicine

Khaled Saab^{○,1}, Tao Tu^{○,‡,1}, Wei-Hung Weng^{○,1}, Ryutaro Tanno^{○,2}, David Stutz^{*,2}, Ellery Wulczyn^{*,1},
Fan Zhang^{*,1}, Tim Strother^{*,1}, Chunjong Park^{*,1}, Elahe Vedadi^{*,1}, Juanma Zambrano Chaves^{*,1},
Sreyas V. V. ^{*,1}, Mihail St. Iovita^{*,1}, Adam W. Kozlowski^{*,2}, Yoon G. Chae^{*,2}, Paul D. M. R. ^{*,2}

Ca
Li
Jo
Si
Pl
Jo
O
D
Al
°C



Multimodal biomedical AI

Julián N. Acosta^{○,1}, Guido J. Falcone¹, Pranav Rajpurkar^{○,2,4} and Eric J. Topol^{○,3,4}

The increasing availability of biomedical data from large biobanks, electronic health records, medical imaging, wearable and ambient biosensors, and the lower cost of genome and microbiome sequencing have set the stage for the development of multimodal artificial intelligence solutions that capture the complexity of human health and disease. In this Review, we outline the key applications enabled, along with the technical and analytical challenges. We explore opportunities in personalized medicine, digital clinical trials, remote monitoring and care, pandemic surveillance, digital twin technology and virtual health assistants. Further, we survey the data, modeling and privacy challenges that must be overcome to realize the full potential of multimodal artificial intelligence in health.

What does multimodal data refer to?

Datasets that incorporate information from multiple sources and modalities, including medical imaging, genetic data, patient-report outcomes and clinical note etc.

Med-Gemini

- ✓ Report (Text)
- ✓ Image
- ✓ Video
- ✓ Signal

Implementation



Hello, I am a primary care physician and this is a chest x-ray for a patient of mine. The formal radiology report is still pending and I would like some help to understand the x-ray. Please write a radiology report for me:

FINDINGS: The heart is at the upper limits of normal size. The mediastinal and hilar contours appear within normal limits. The lungs appear clear. There are no pleural effusions or pneumothorax. Mild degenerative changes are noted along the lower thoracic spine. **IMPRESSION:** No evidence of acute disease.



Prompt



...



...



You are a helpful medical video assistant.

Task:

Assess whether Critical View of Safety (CVS) is fully achieved in the provided cholecystectomy video.

The Critical View of Safety (CVS) is fully achieved if the following three criteria are met:

- C1: Clear view of 2 tubular structures connected to the gallbladder.
- C2: A carefully dissected hepatocystic triangle presenting an unimpeded view of only the 2 cystic structures and the cystic plate.
- C3: The lower third of the gallbladder is dissected off the cystic plate.

Instructions:

1. Watch the video carefully, and answer which of the Critical View of Safety (CVS) criteria are met.
2. Output Format:
 - CVS Assessment:
 - C1 - achieved / partially achieved / not achieved (Optional explanation)
 - C2 - achieved / partially achieved / not achieved (Optional explanation)
 - C3 - achieved / partially achieved / not achieved (Optional explanation)



Response

CVS Assessment:

C1 - partially achieved: Two tubular structures are visible but the view is not entirely clear due to some bleeding obscuring the area.

C2 - not achieved: The hepatocystic triangle is not clearly dissected, and there is not an unimpeded view of only the two cystic structures and the cystic plate.

C3 - not achieved: The lower third of the gallbladder is not dissected off the cystic plate.

Multimodal Data

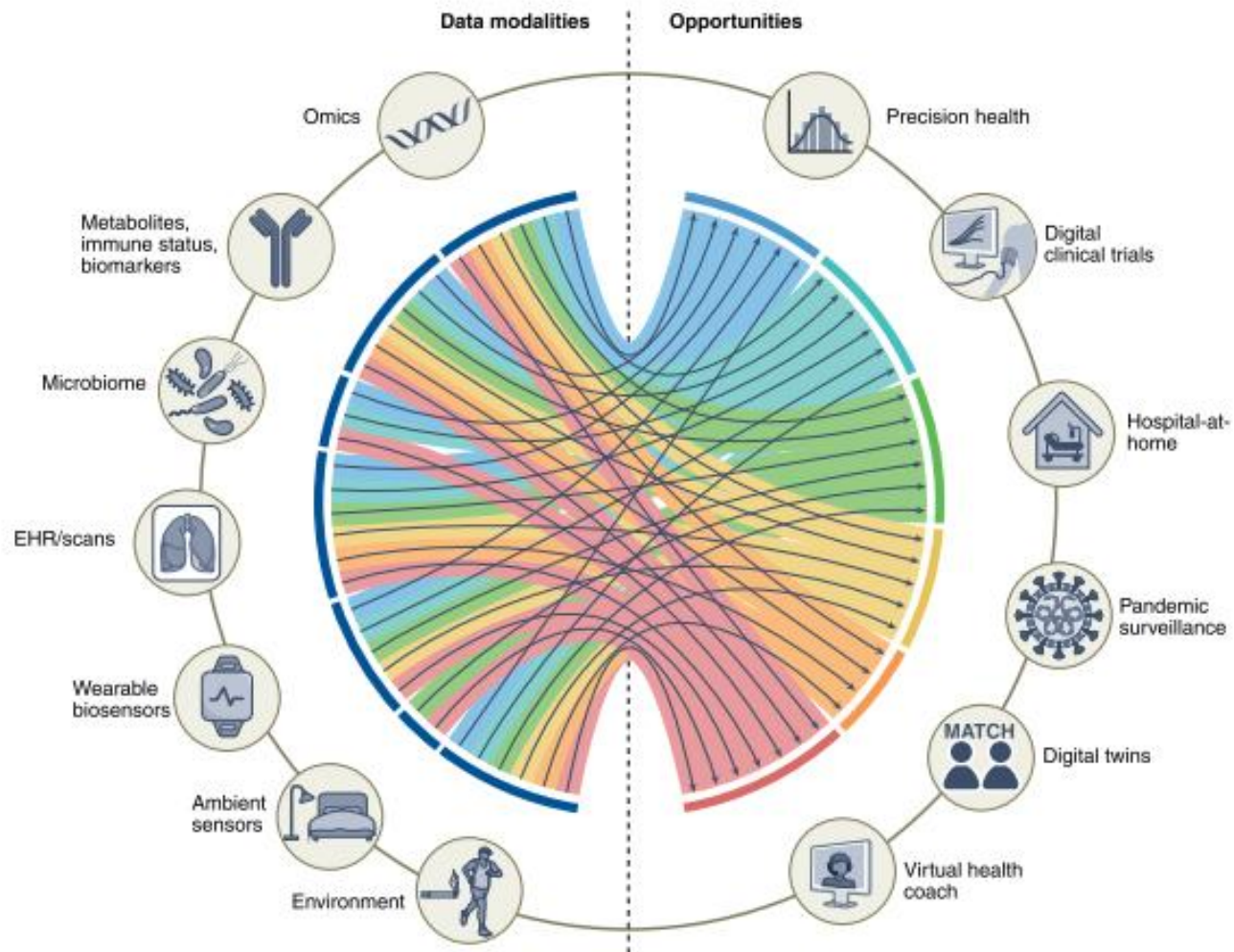
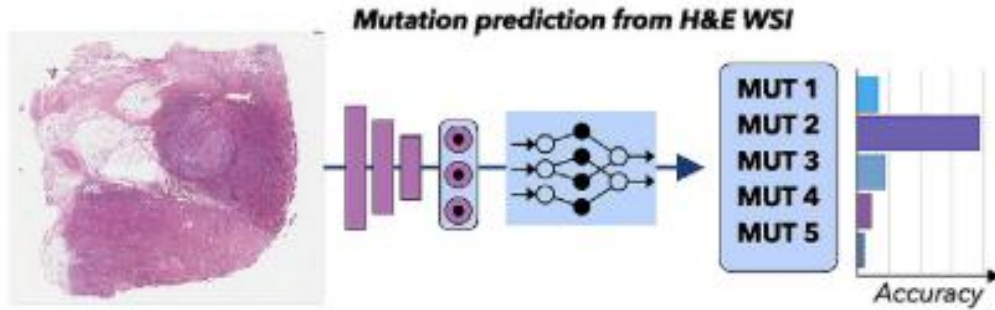


Fig. 1 | Data modalities and opportunities for multimodal biomedical AI. Created with BioRender.com.

- **EHR**
 - ✓ imaging
 - ✓ patient-report outcome
 - ✓ clinical notes
- **Wearable Biosensor**
 - ✓ EEG
 - ✓ ECG
 - ✓ EMG
- **Ambient Sensor**
 - ✓ video or depth cameras;
 - ✓ microphones
- **Omics**
 - ✓ genetic data (WGS)
 - ✓ transcriptome
 - ✓ proteome
 - ✓ metabolome
 - ✓ antibody

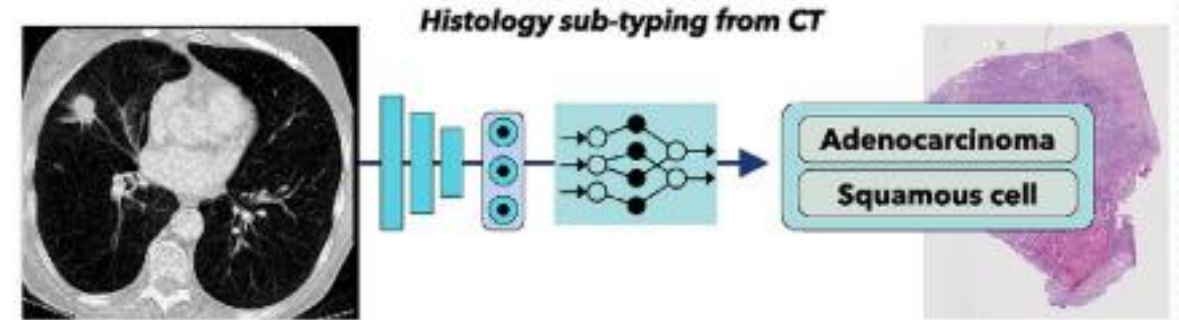
Implementation

A MORPHOLOGIC ASSOCIATES



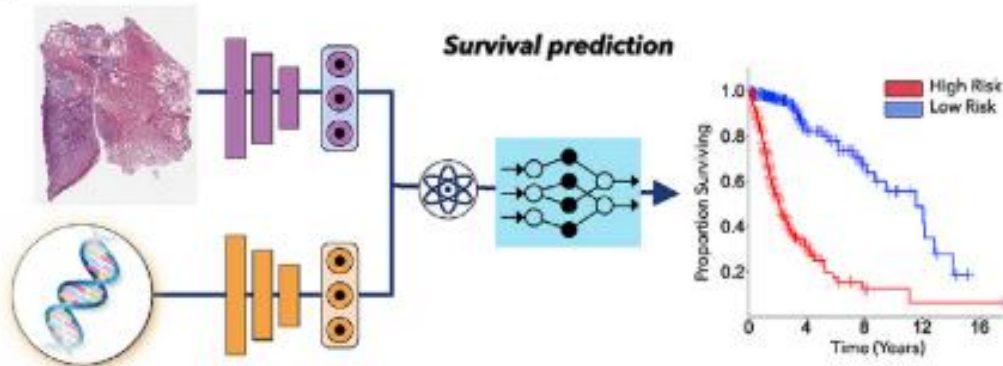
(A) Infer certain mutations from histology

B NON-INVASIVE ALTERNATIVES



(B) Find the relationship between non-invasive and invasive modalities, such as prediction of histology subtype from radiomics features.

C OUTCOME ASSOCIATES



(C) Uncover associations between clinical data and patient outcome, contributing to the discovery of predictive features within and across modalities.

D EARLY PREDICTORS

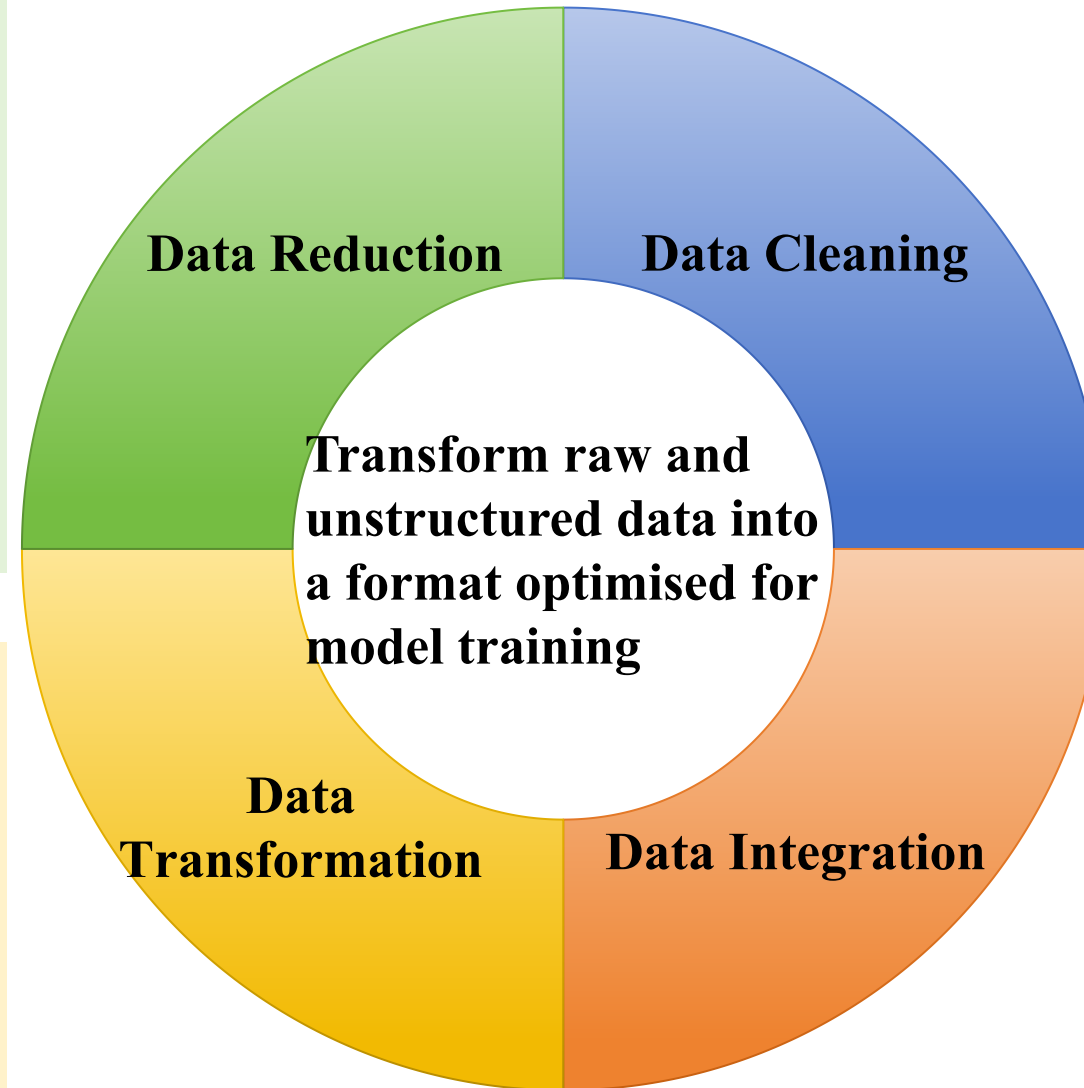


(D) Identify risk factors related to cancer onset or uncover patterns related with treatment response or resistance, to support early interventions.

Multimodal Data Pre-processing

- **Feature selection:**
 - ✓ Region of interest in medical imaging
 - ✓ select associated genetic sequence
 - ✓ select features from laboratory test
- **Dimensionality reduction:**
 - ✓ reduce the volume (3D --> 2D)

- **Raw text data:**
 - ✓ Vector embeddings
- **Genetic sequence:**
 - ✓ Vector embedding
- **Medical Imaging :**
 - ✓ re-scaling
 - ✓ data augmentation

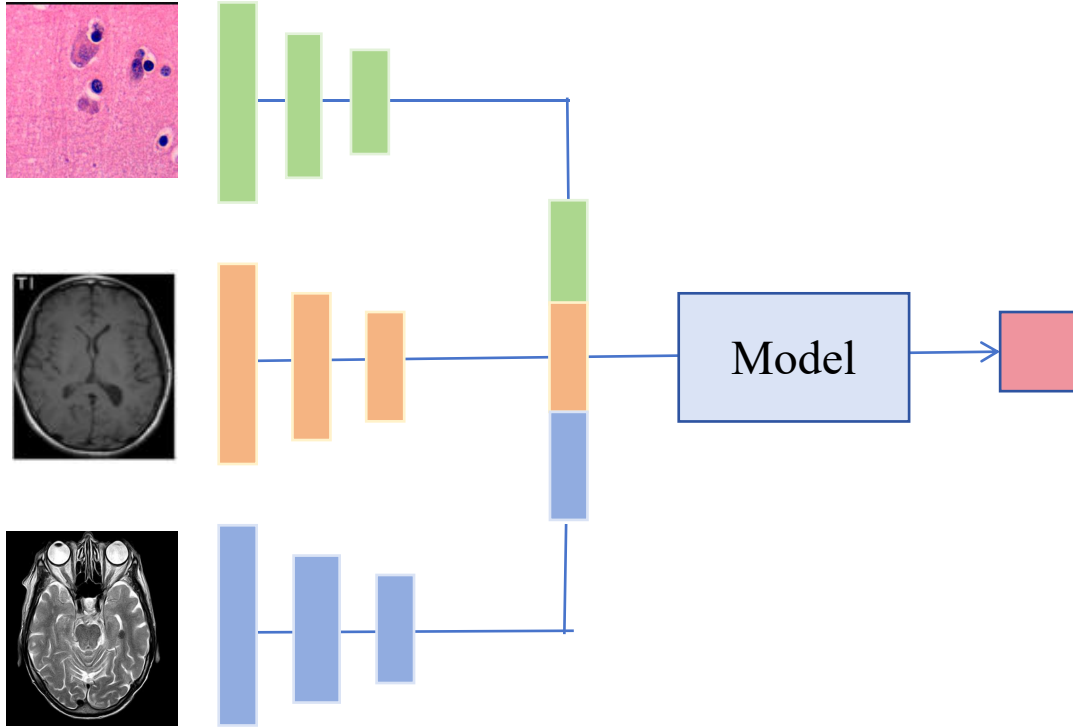


- **Impute Missing Values:**
 - ✓ laboratory test: the test was not ordered; insufficient for analysis
 - ✓ vital signs: temperature; blood pressure; heart rate
- **Remove Duplicates from the Raw Dataset:**
 - ✓ duplicate records for the same patient (system issue)

- Create an Unified Dataset Merged:**
- ✓ Medical Imaging
 - ✓ Disease associated genes
 - ✓ Laboratory Test

Multimodal Data Fusion

Early Fusion

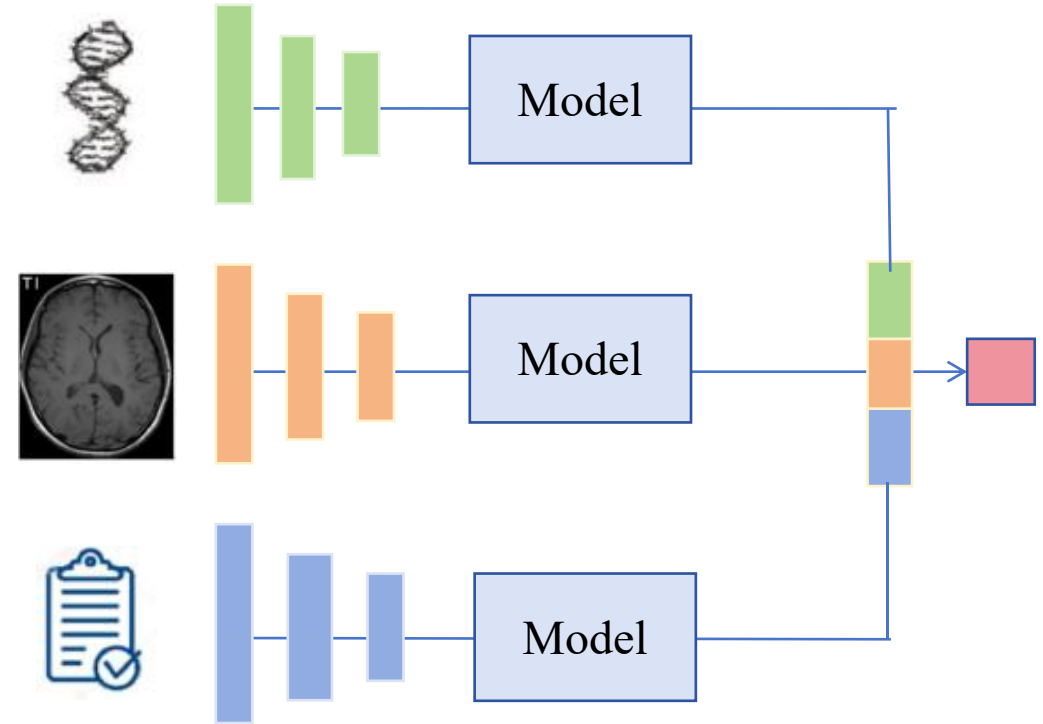


Definition: refer to joint representation from raw data or features at the input level before feeding it to the model.

Keypoint: integration of similar modalities.

Feature Fusion: concatenation

Late Fusion



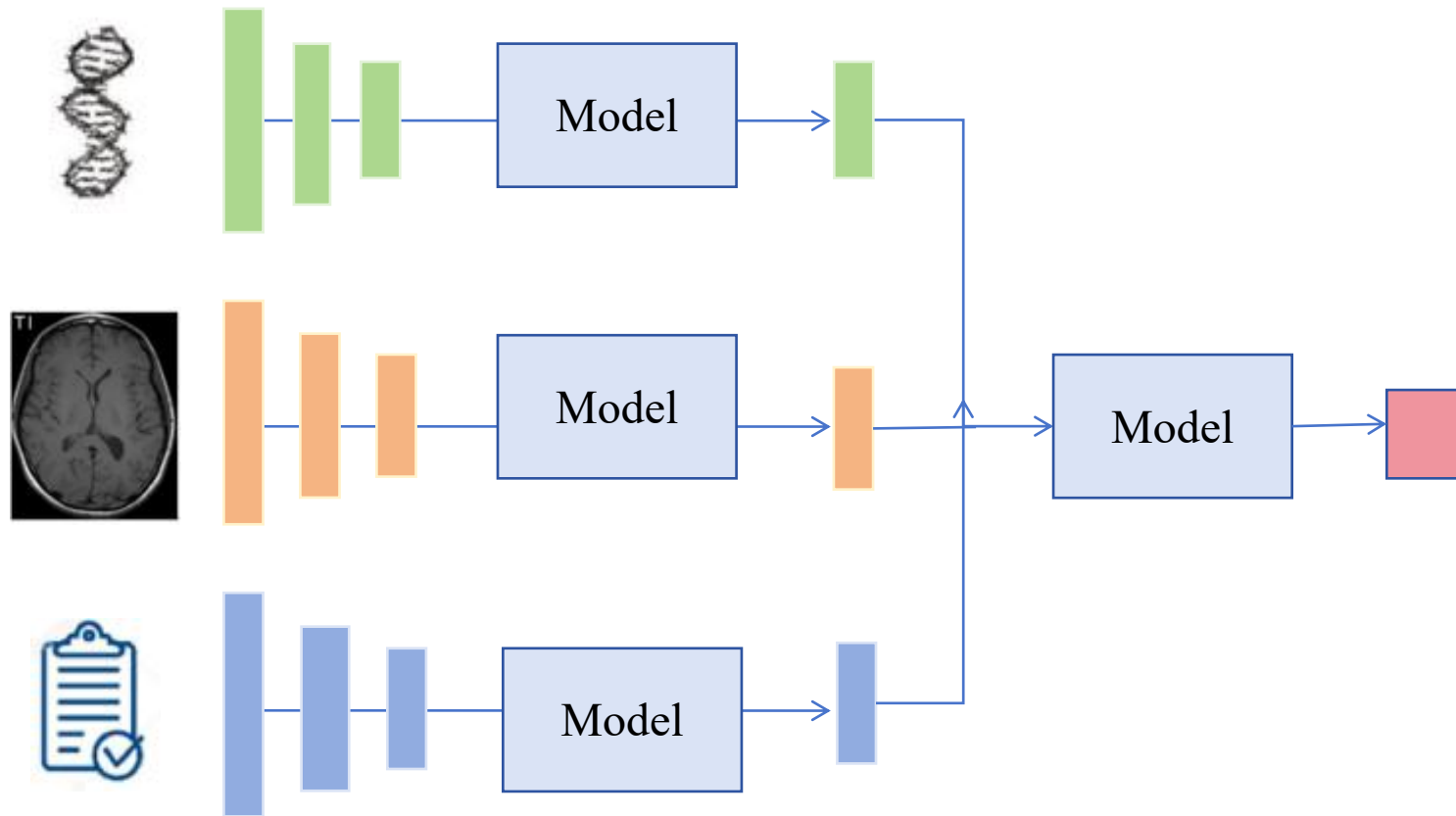
Definition: refer to decision-level fusion, which trains a separate model for each modality and aggregates the predictions from individual models for the final prediction.

Keypoints: heterogeneous data; without data synchronization limits

Feature Fusion: averaging, majority voting, Bayes-based rules

Multimodal Data Fusion

Intermediate Fusion



Definition

The prediction can supervise the feature extraction models to improve feature representations under the multimodal context

Keypoints

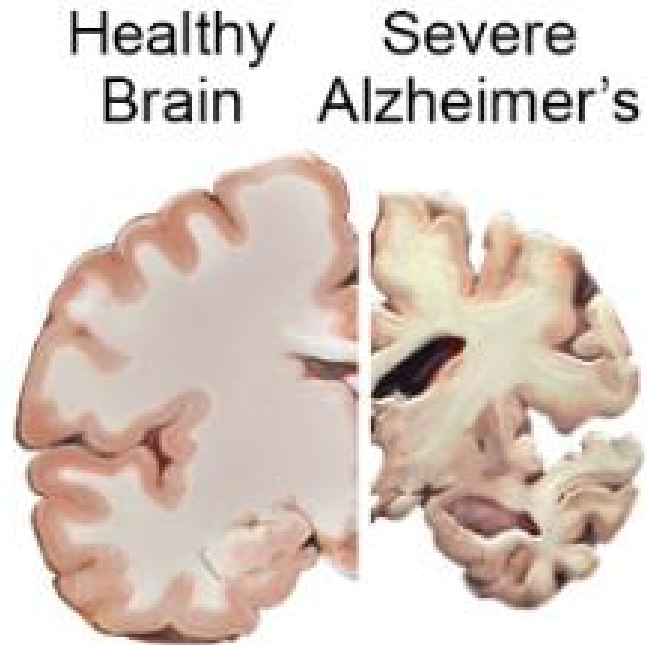
- unimodal embedding can be learned under the multimodal context
- Don't have data synchronization limits
- Tolerate heterogeneous data

Feature Fusion

- Attention Mechanism
- Graph Neural Network

Alzheimer's Disease Detection

Alzheimer's disease (AD) is a progressive neurologic disorder that affects more than 850,000 people in the UK, costing the UK economy around £26bn per year.



How is the brain affected:

1. The damage initially appears to take place in the **hippocampus**.
2. Then additional parts of the brain are affected and begin to **shrink**.
3. Finally, damage is widespread and brain tissue has shrunk significantly.

Diagnosis:

Reliable diagnoses can be confirmed by complementary information among **multi-modalities** such as images, cerebrospinal fluid, and blood test ect.

Challenge: It is difficult for experts to recognize all shifts in all modalities for AD patients, leading to misdiagnosis.

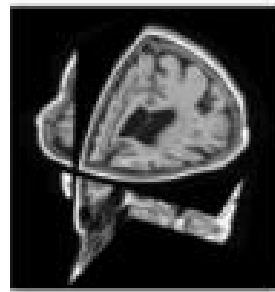
Alzheimer's Disease Detection

Aims

Create an Automatic Explainable Alzheimer Disease Detection Framework Based on Multi-Modal Data

Objectives

- Precisely extract and combine features from each modality including imaging, genetic and tabular data
- Figure out why the automatic framework makes such a decision, making the detection result trustworthy



Alzheimer's Disease Detection

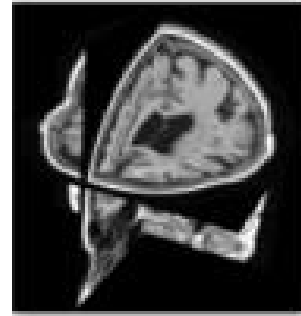
Data Types

Tabular Data



- Structured clinical data (laboratory test data)
- neurological exams (balance test)
- cognitive assessments (memory tests)

Imaging Data



- Select slices from the 3D MRI

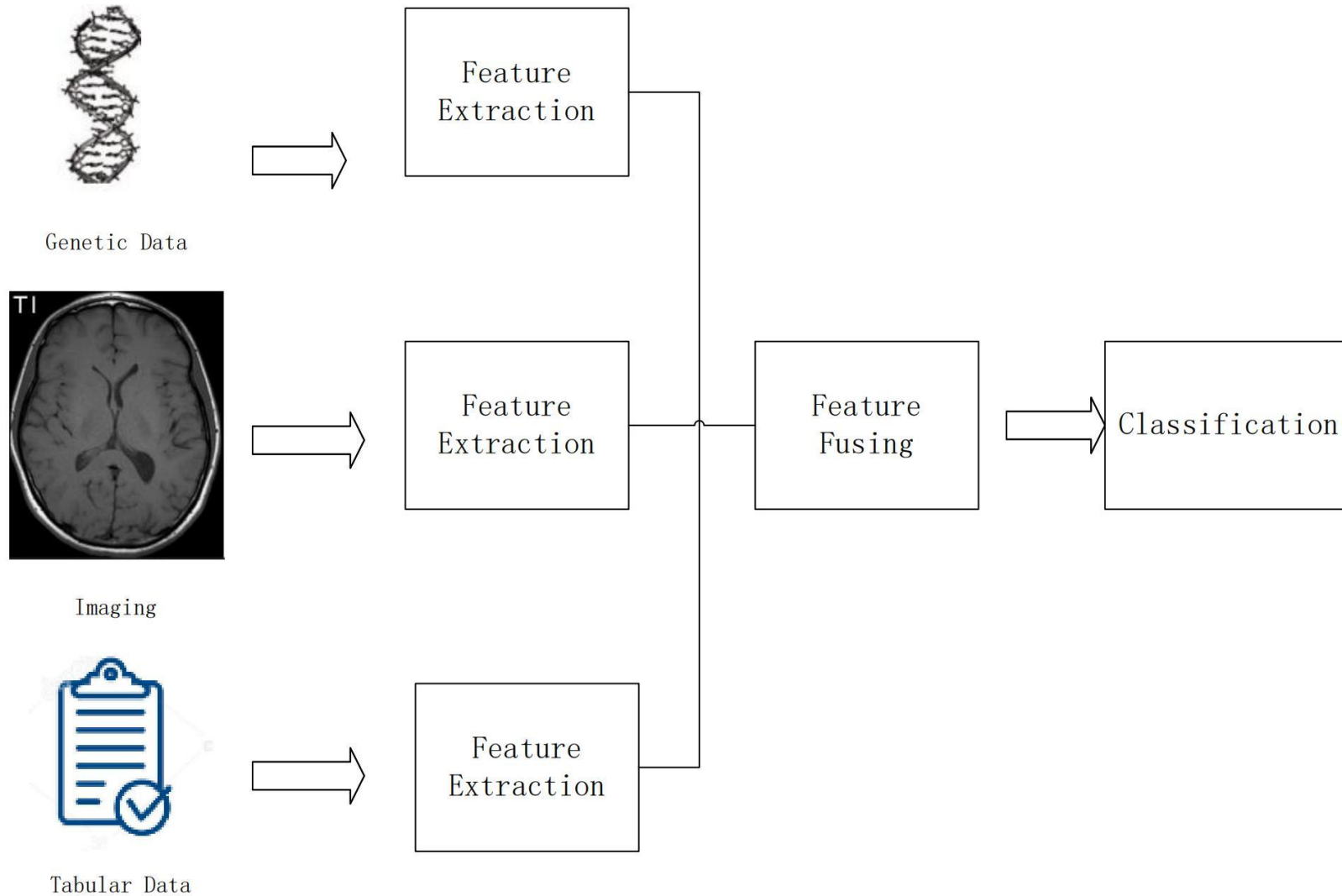
Genetic Data



SNPs selected by the genes of known association with AD (none allele; both allele and; one allele)

Alzheimer's Disease Detection

Framework



Diagnosis Result:

- Normal Control
- Early Mild Cognitive Impairment
- Late Cognitive Impairment
- Alzheimer's Disease

Alzheimer's Disease Detection

Challenges

- The **feature association** among different modalities should be explored.
- **Feature significance** for each modality should be explored.
- Generate **human-understandable** explanation of the decision made by the machine.

Thanks for Listening!

Email: yuhan.4.wang@kcl.ac.uk