La Salle College

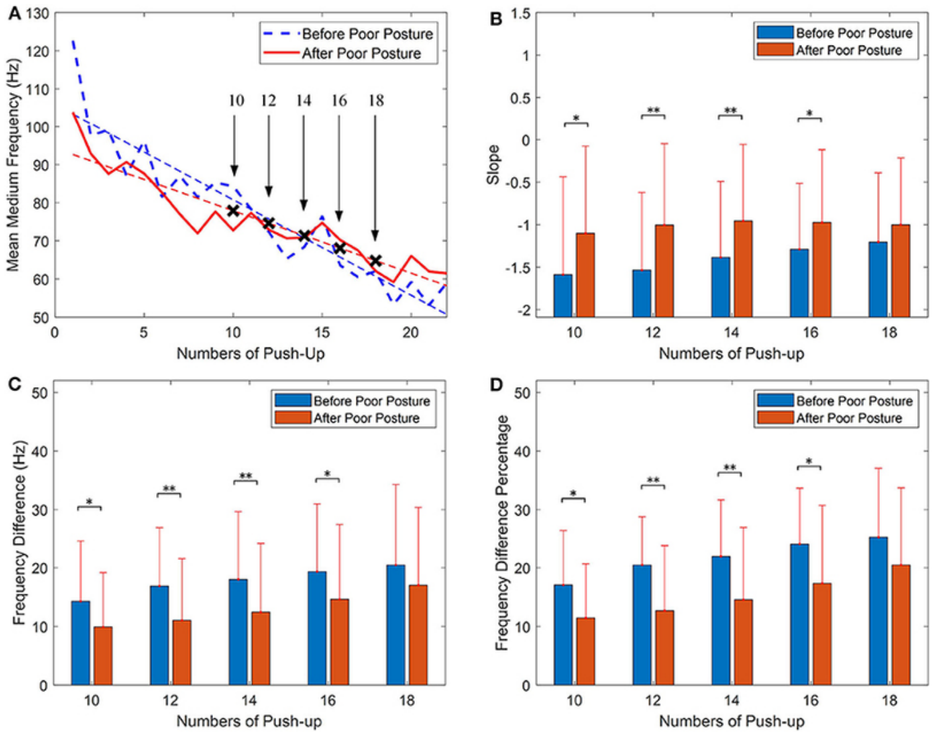
PosturePal

姿勢好幫手

Eason Xie, Michael Lam, Jason Shang

Why should we sit correctly?

- **Chronic back** and **neck pain**
- Increased risk of **musculoskeletal disorders**
- **Stress**, **reduced self-esteem**, and productivity
- Even **cancer** or **heart diseases** in the long term*



*Department of Health & Human Services. (n.d.). The dangers of sitting: why sitting is the new smoking. Better Health Channel. https://www.betterhealth.vic.gov.au/health/healthyliving/the-dangers-of-sitting

Status Quo

- **Poor awareness** of posture habits
- **Lack of accessible tools** for correction
- **Low motivation** due to work/study pressures, especially **students**
- Cycle of neglect worsening symptoms

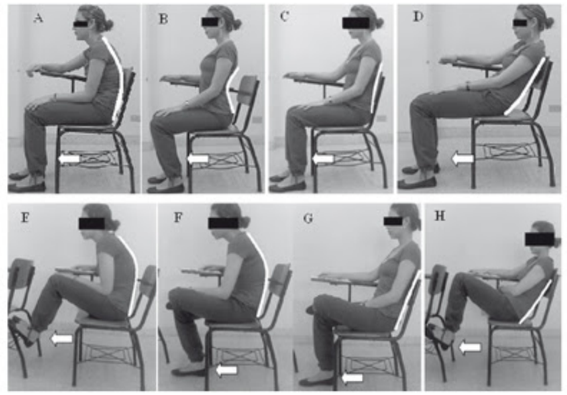


Figure 1. Sitting posture categories (A) Rounded back or increase of the kyphosis with the feet supported on the floor; (B) Increase of the lordosis and the feet supported on the

	Mean ± SD	Minimum-Maximum.
BackPEI posture score	1.42 ± 0.98	0-5
	n	%
Sitting posture while listening teacher on desk		
Correct body posture	312	14.0
Poor posture	1,909	86.0
Sitting posture while writing on desk		
Correct body posture	75	3.4
Poor posture	2,146	96.6
Sitting posture while using computer		
Correct body posture	338	15.2
Poor posture	1,883	84.8
Posture while grabbing something on the floor		
Correct body posture	928	41.8
Poor posture	1,293	58.2
Posture while carrying a school backpacks* (n = 2,036)		
Correct body posture	1,448	71.1
Poor posture	588	28.9

SD = standard deviation.
* Adolescents who carry a school backpacks.

Existing Solutions

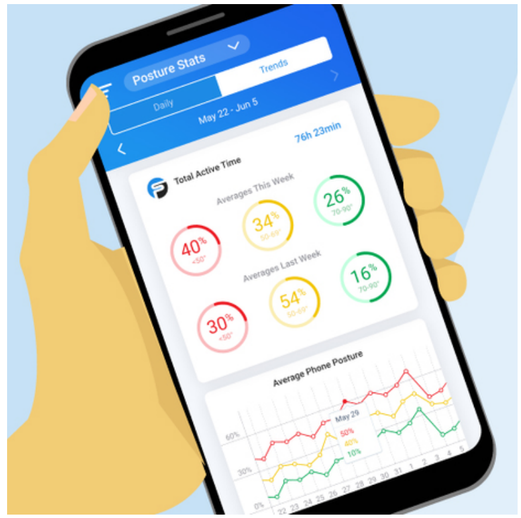
Ergonomic furnitures

High cost
Passive support

Wearable devices

Diminishing effectiveness
Comfort issues

Posture Monitoring Apps

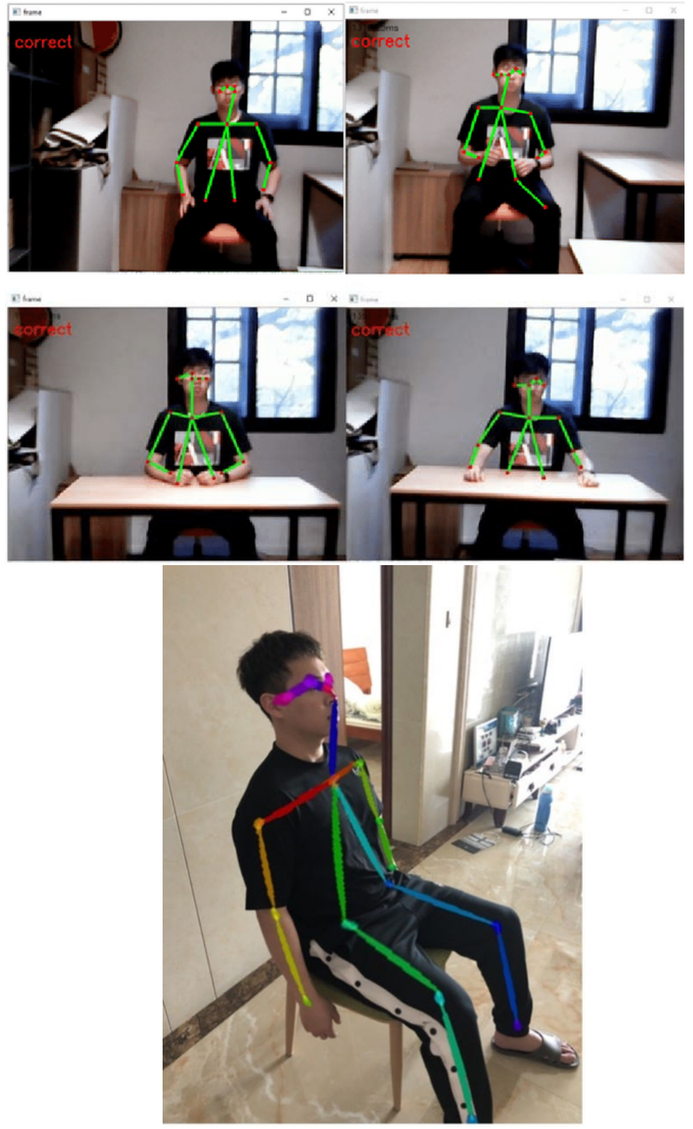


Generic feedback
Engagement challenges

Our Solution

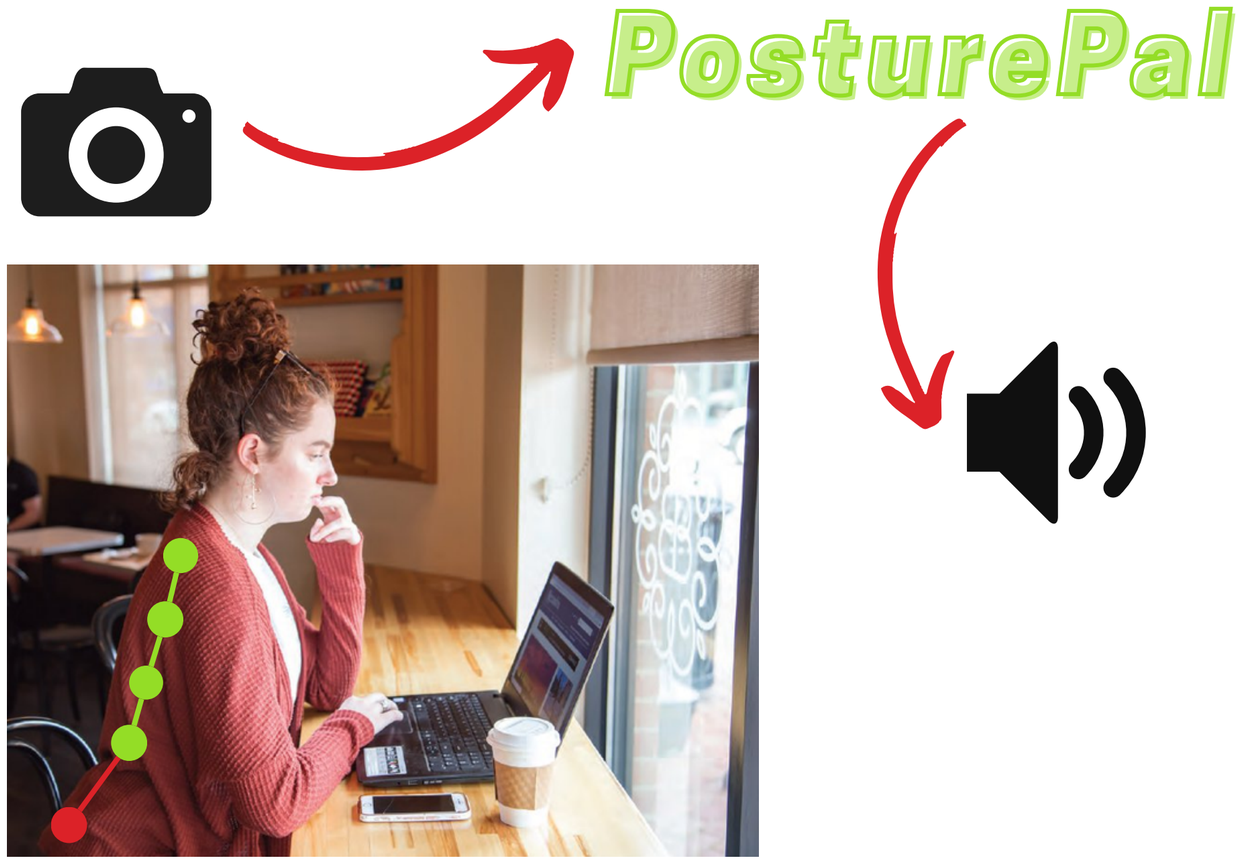
PosturePal

- **AI**-driven
- **Sensors + AI algorithms + Large Language Models (LLM)**
- **Humorous** Feedback
- Focuses on **“Fun”**



Advantages of PosturePal

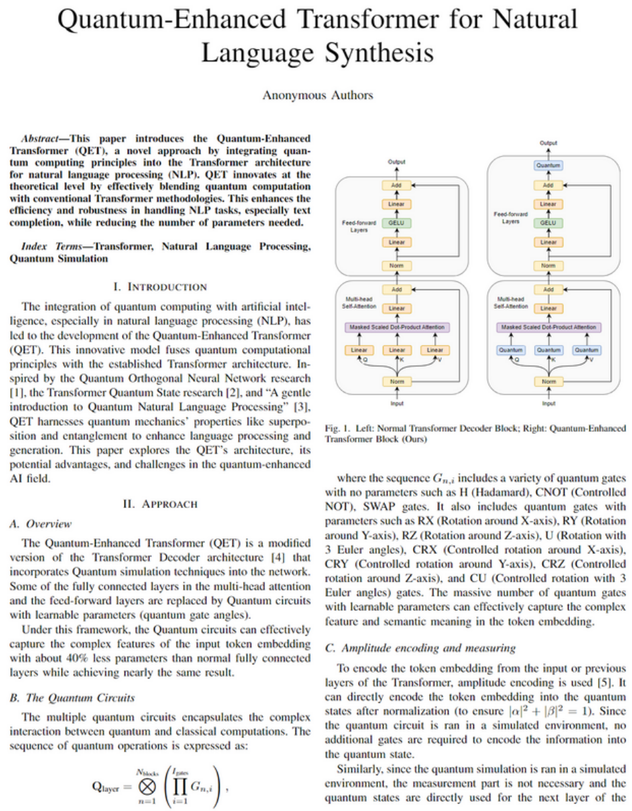
- Personalization
- User Engagement
- Sustainability
- Accessibility
- Feedback Mechanism



Prototype Development

Some of our previous works

- **Version 1** of PosturePal: Implemented on Scratch (5 years ago)
- **Version 2** of PosturePal: PoseNet-only, no LLM or custom ResNet model (3 years ago)
- **Published a paper on IEEE CAI 2024** of Quantum-Enhanced Transformer: Integrating quantum computing techniques into the Transformer architecture for improved capability & decreased parameters.



Quantum-Enhanced Transformer for Natural Language Synthesis

Anonymous Authors

***Abstract*—This paper introduces the Quantum-Enhanced Transformer (QET), a novel approach by integrating quantum computing principles into the Transformer architecture for natural language processing (NLP). QET innovates at the theoretical level by effectively blending quantum computation with conventional Transformer methodologies. This enhances the efficiency and robustness in handling NLP tasks, especially text completion, while reducing the number of parameters needed.**

***Index Terms*—Transformer, Natural Language Processing, Quantum Simulation**

I. Introduction

The integration of quantum computing with artificial intelligence, especially in natural language processing (NLP), has led to the development of the Quantum-Enhanced Transformer (QET). This innovative model fuses quantum computational principles with the established Transformer architecture. Inspired by the Quantum Orthogonal Neural Network research [1], the Transformer Quantum State research [2], and "A gentle introduction to Quantum Natural Language Processing" [3], QET harnesses quantum mechanics' properties like superposition and entanglement to enhance language processing and generation. This paper explores the QET's architecture, its potential advantages, and challenges in the quantum-enhanced AI field.

II. Approach

A. Overview

The Quantum-Enhanced Transformer (QET) is a modified version of the Transformer Decoder architecture [4] that incorporates Quantum simulation techniques into the network. Some of the fully connected layers in the multi-head attention and the feed-forward layers are replaced by Quantum circuits with learnable parameters (quantum gate angles).

Under this framework, the Quantum circuits can effectively capture the complex features of the input token embedding with about 40% less parameters than normal fully connected layers while achieving nearly the same result.

B. The Quantum Circuits

The multiple quantum circuits encapsulates the complex interaction between quantum and classical computations. The sequence of quantum operations is expressed as:

$$\mathbf{Q}_{\text{layer}} = \bigotimes_{n=1}^{N_{\text{blocks}}} \left(\prod_{i=1}^{t_{\text{gates}}} G_{n,i} \right),$$

Fig. 1. Left: Normal Transformer Decoder Block; Right: Quantum-Enhanced Transformer Block (Ours)

where the sequence $G_{n,i}$ includes a variety of quantum gates with no parameters such as H (Hadamard), CNOT (Controlled NOT), SWAP gates. It also includes quantum gates with parameters such as RX (Rotation around X-axis), RY (Rotation around Y-axis), RZ (Rotation around Z-axis), U (Rotation with 3 Euler angles), CRX (Controlled rotation around X-axis), CRY (Controlled rotation around Y-axis), CRZ (Controlled rotation around Z-axis), and CU (Controlled rotation with 3 Euler angles) gates. The massive number of quantum gates with learnable parameters can effectively capture the complex feature and semantic meaning in the token embedding.

C. Amplitude encoding and measuring

To encode the token embedding from the input or previous layers of the Transformer, amplitude encoding is used [5]. It can directly encode the token embedding into the quantum states after normalization (to ensure $|\alpha|^2 + |\beta|^2 = 1$). Since the quantum circuit is ran in a simulated environment, no additional gates are required to encode the information into the quantum state.

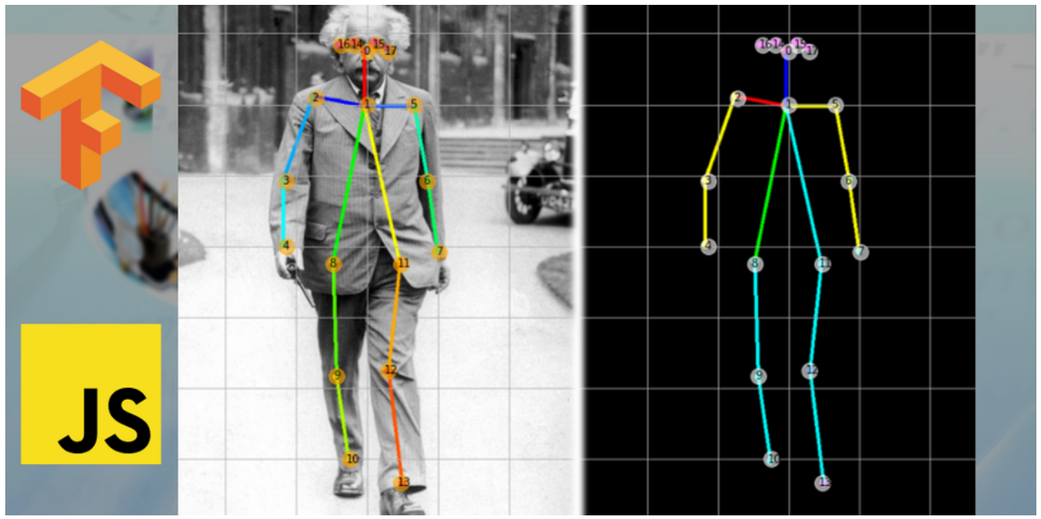
Similarly, since the quantum simulation is ran in a simulated environment, the measurement part is not necessary and the quantum states are directly used for the next layer of the

How It Works

- Detect the user’s posture with a **finetuned PoseNet**
- Analysation with our **custom model**
- Send the analysed results to the LLM for **personalized & fun reminders**
- Read the LLM’s results with text-to-speech (TTS)

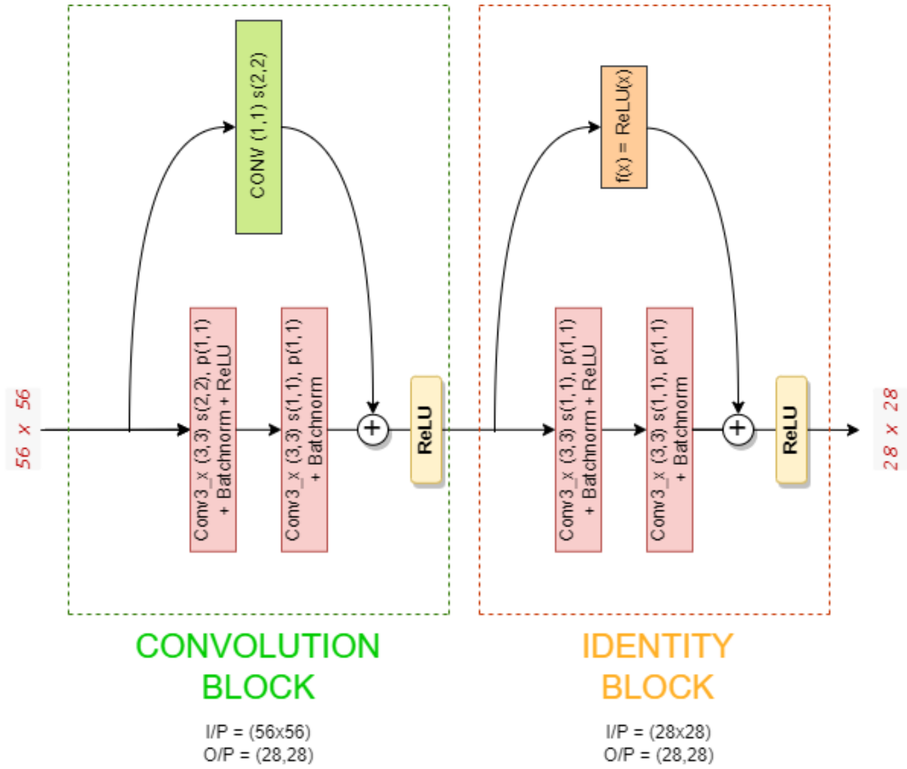
PoseNet

- A **MobileNet** model trained by google with **Tensorflow Lite**
- Detects the **key body points**
- We converted it to **PyTorch** format
- We **finetuned** it with custom upper-body data for **improved accuracy**



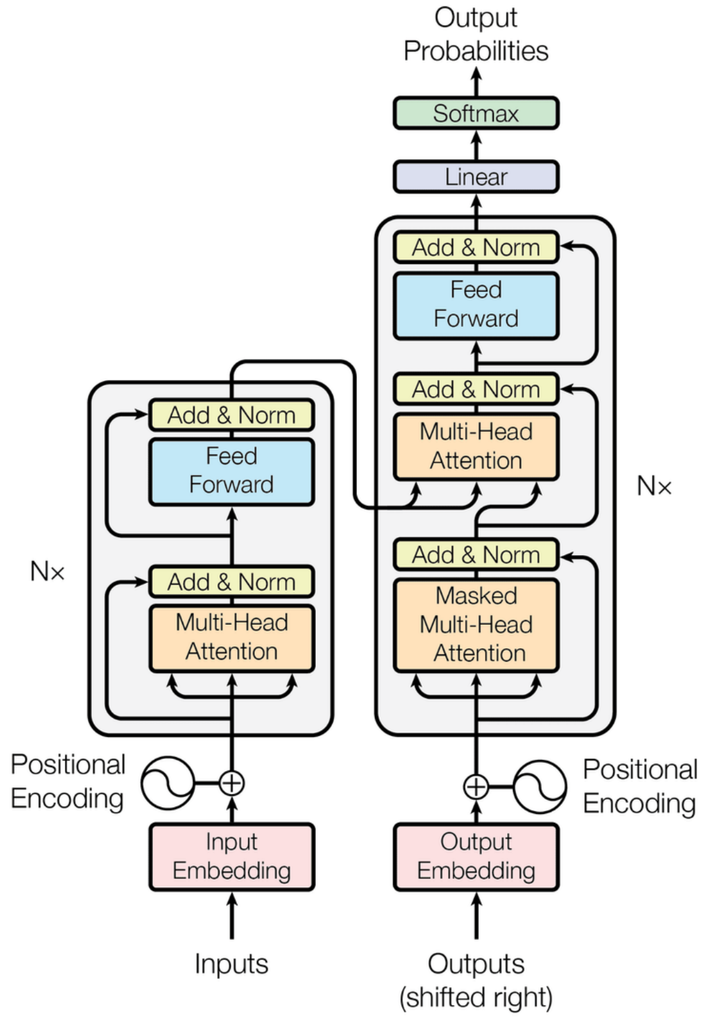
Custom Model for analysation

- A custom **ResNet** architecture **Convolutional Neural Network (CNN)**
- Trained on **custom-made data**
- Takes the output of PoseNet and detect whether the user is sitting correctly or not



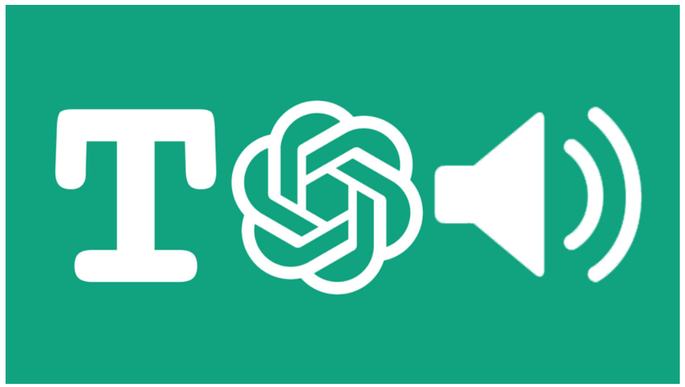
LLM for fun reminder

- Mistral 8x7B Large Language Model
- **Finetuned** with recommended responses
- Generates **fun & personalized** encouragements / reminders for the user



OpenAI Text-to-Speech API

- Reads out the generated texts by the LLM
- An “Audio User Interface”
- Users can choose between 6 different voices
- Gives the user a **friendly & nice feel**



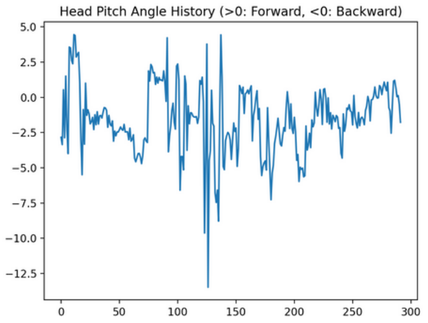
Prototype Demostration

Good Sitting Posture

Bad Sitting Posture

Prototype Testing

- We invited some of our schoolmates to test our prototype
- Our prototype received an average score of **8.6/10**
- Most of them **liked our system & think it can help them** to correct their sitting postures

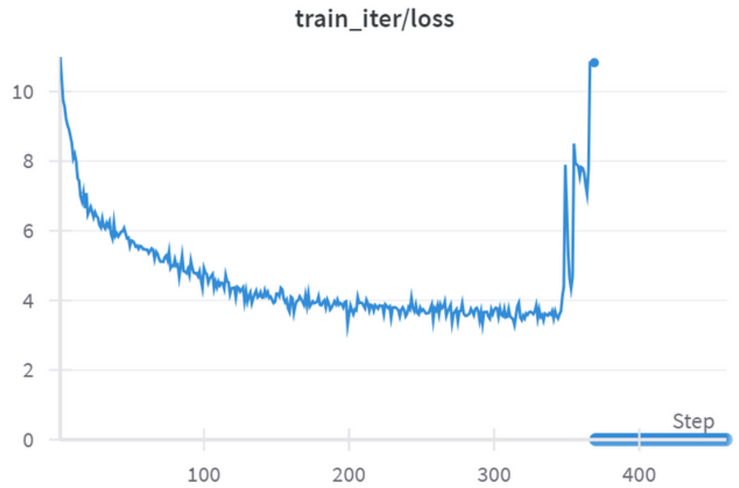


Difficulties & Challenges

- Real-time processing and analysis latency
- **Suboptimal hyperparameters** for training and finetuning
- Custom ResNet model **overfitting**
- Dynamic & fun LLM response generation

```
Traceback (most recent call last):
  File "C:\Users\eason\OneDrive - La Salle College\vs_code_2022\AI\PosturePal\run.py", line 221, in <module>
    main()
  File "C:\Users\eason\OneDrive - La Salle College\vs_code_2022\AI\PosturePal\run.py", line 174, in main
    roll, pitch = resnet(input_image, keypoint_coord)
  File "C:\Users\eason\AppData\Local\Programs\Python\Python39\lib\site-packages\torch\nn\modules\module.py", line 1511,
in _wrapped_call_impl
    return self._call_impl(*args, **kwargs)
  File "C:\Users\eason\AppData\Local\Programs\Python\Python39\lib\site-packages\torch\nn\modules\module.py", line 1520,
in _call_impl
    return forward_call(*args, **kwargs)
  File "C:\Users\eason\OneDrive - La Salle College\vs_code_2022\AI\PosturePal\posenet\resnet.py", line 35, in forward
    roll, pitch = self.model(image)
  File "C:\Users\eason\OneDrive - La Salle College\vs_code_2022\AI\PosturePal\posenet\resnet.py", line 30, in model
    pitch_angle = image[1].max()
IndexError: index 1 is out of bounds for dimension 0 with size 1
```

```
Average FPS:  2.7269552570218814
```



Future Works

- Improve the accuracy of sitting posture detection with **MoveNet** (Goyal et. al., 2023) or even based on **pressure distribution** (Seo et. al., 2021)
- Decrease the computational complexity for running on low-spec laptops / tablets 💻
- Release to the public for more feedbacks to improve the system
- More advanced feedback systems, e.g. mini games, better jokes 🤣

MoveEnet: Online High-Frequency Human Pose Estimation with an Event Camera

Gaurvi Goyal, Franco Di Pietro, Nicolo Carissimi, Arren Glover, Chiara Bartolozzi
Event-Driven Perception for Robotics
Istituto Italiano di Tecnologia, Italy
{gaurvi.goyal, franco.dipietro, nicolo.carissimi, arren.glover, chiara.bartolozzi}@iit.it

Abstract

Human Pose Estimation (HPE) is crucial as a building block for tasks that are based on the accurate understanding of human position, pose and movements. Therefore, accuracy and efficiency in this block echo throughout a system, making it important to find efficient methods, that run at fast rates for online applications. The state of the art for mainstream sensors has made considerable advances, but event camera based HPE is still in its infancy. Event cameras boast high rates of data capture in a compact data structure, with advantages like high dynamic range and low power consumption. In this work, we present a system for a high frequency estimation of 2D, single-person Human Pose with event cameras. We provide an online system, that

Figure 1. Sample result of MoveEnet from the event-Human 3.6m dataset superimposed on EROS representation

cameras has increased in the recent years, as they deliver information in a more compact data format, thus,

Implementation of CNN Model for Classification of User Sitting Posture Based on Pressure Distribution

Ji-Yun Seo, Ji-Su Lee, Sang-Joong Jung, Yun-Hong Noh & Do-Un Jeong

Conference paper | First Online: 06 February 2021

1076 Accesses

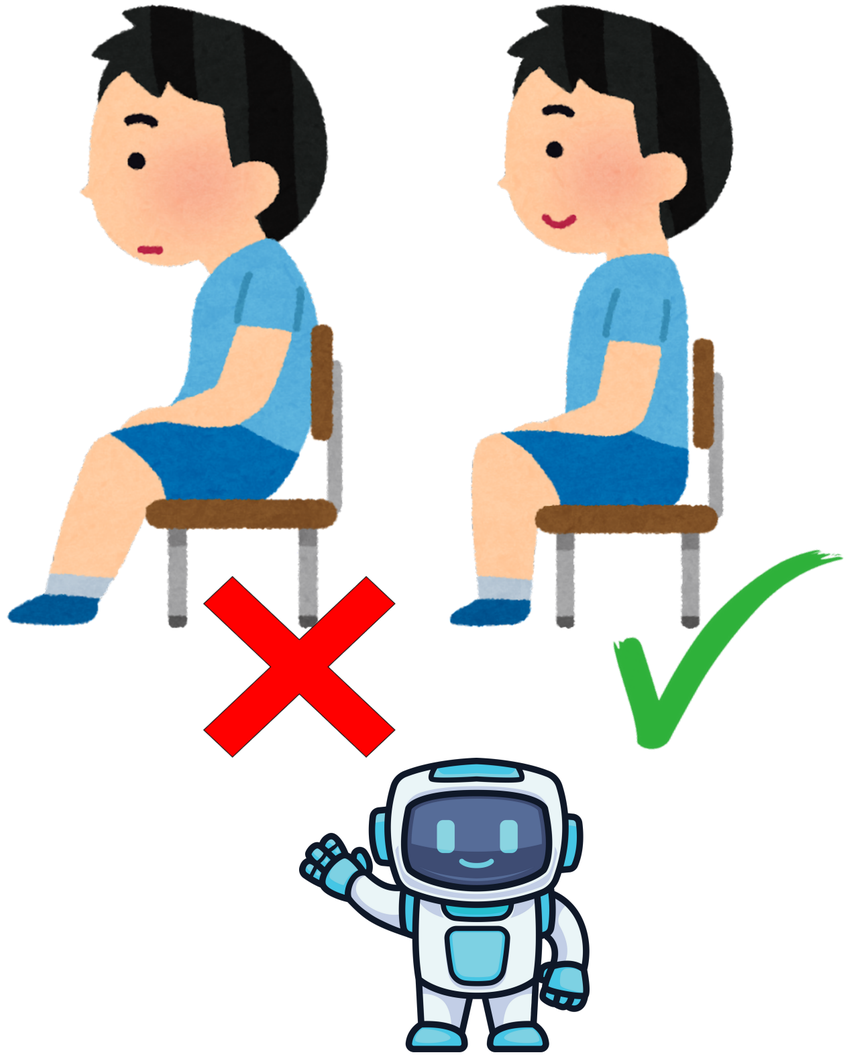
Part of the Lecture Notes in Computer Science book series (LNISA,volume 12616)

Abstract

Musculoskeletal disease is often caused by sitting down for long period's time or by bad posture habits. In order to prevent musculoskeletal disease in daily life, it is the most important to correct the bad sitting posture to the right one through real-time monitoring. In this study, to detect the sitting information of user's without any constraints, we propose

Conclusion

- Solve the problem of **incorrect sitting posture** with a **fun & personalized** approach
- Utilizes **LLMs** for **engaging, interactive feedback** on posture
- Tailors feedback to resonate with individual **user interests** and **humor**
- Innovates **human-machine communication** with fun, interactive reminders



La Salle College

Thank you