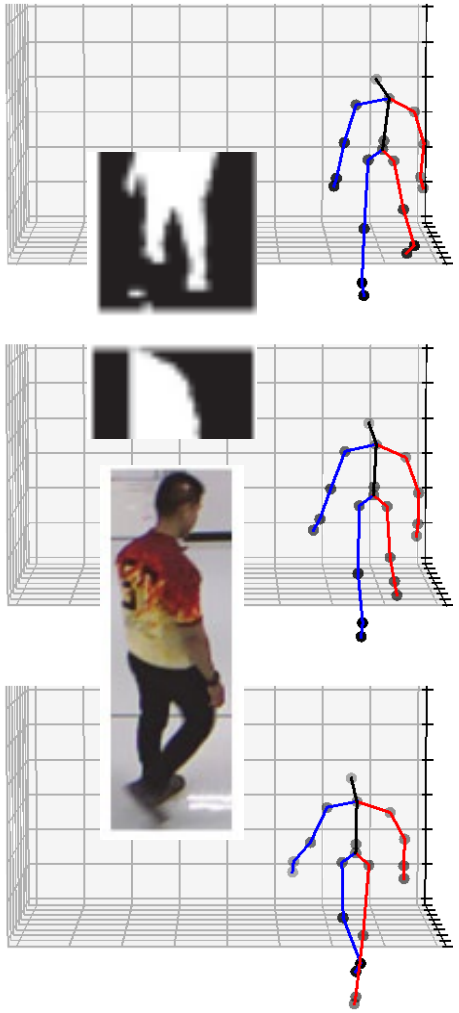
Gait Recognition using Machine Learning Techniques



รองศาสตราจารย์ ดร. นิรัตยา คำเสามานันทน์

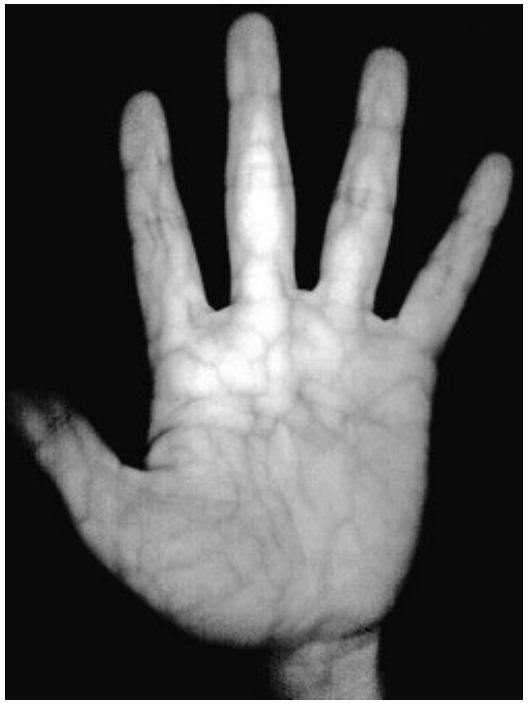
หน่วยวิจัยด้านการวิเคราะห์ท่าทางการเคลื่อนไหวและเทคโนโลยีชาญฉลาด (GaitTech)

สถาบันเทคโนโลยีนานาชาติสิรินธร มหาวิทยาลัยธรรมศาสตร์

(งานวิจัยนี้ดำเนินการร่วมกับ รศ. ดร. ชลวิช นัทธี และนักศึกษาป.โท/เอก)

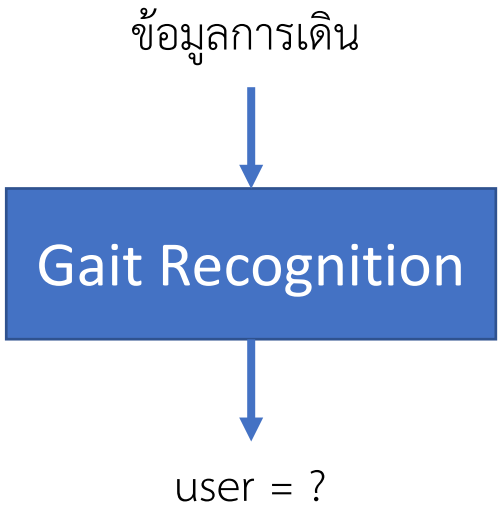
Biometric Authentication

- Biometric Authentication เป็นวิธีการใช้ข้อมูลชีวมาตรแบบต่าง ๆ เพื่อพิสูจน์/ยืนยันตัวบุคคล ตัวอย่างเช่น
  - ลายนิ้วมือ
  - ม่านตา
  - ลายเส้นเลือดบนฝ่ามือ
  - ใบหน้า



Gait Recognition

- Gait Recognition เป็นการศึกษาเพื่อพัฒนาเทคนิคยืนยันตัวบุคคลโดยใช้ท่าทางการเดิน โดยมีจุดเด่น ได้แก่
  - เก็บข้อมูลท่าทางการเดินได้จากระยะไกล
  - ไม่รบกวนบุคคลที่ถูกเก็บข้อมูล
  - ยากที่จะเปลี่ยนแปลงท่าทางการเดินตลอดเวลา

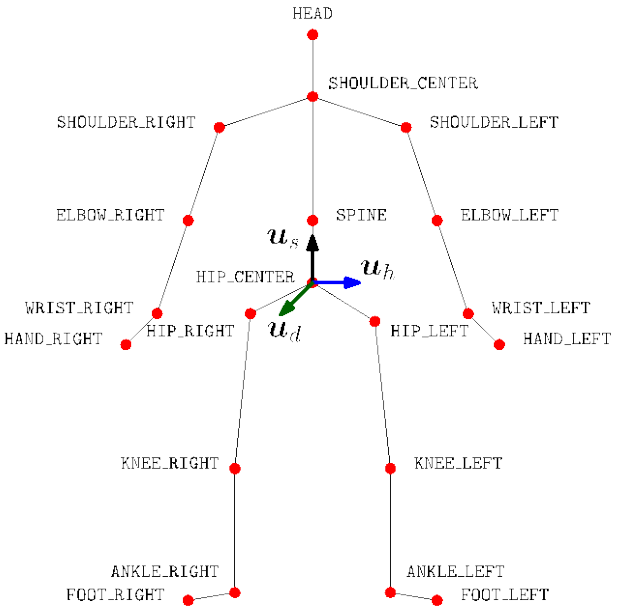


Gait Recognition Techniques

Model-free techniques

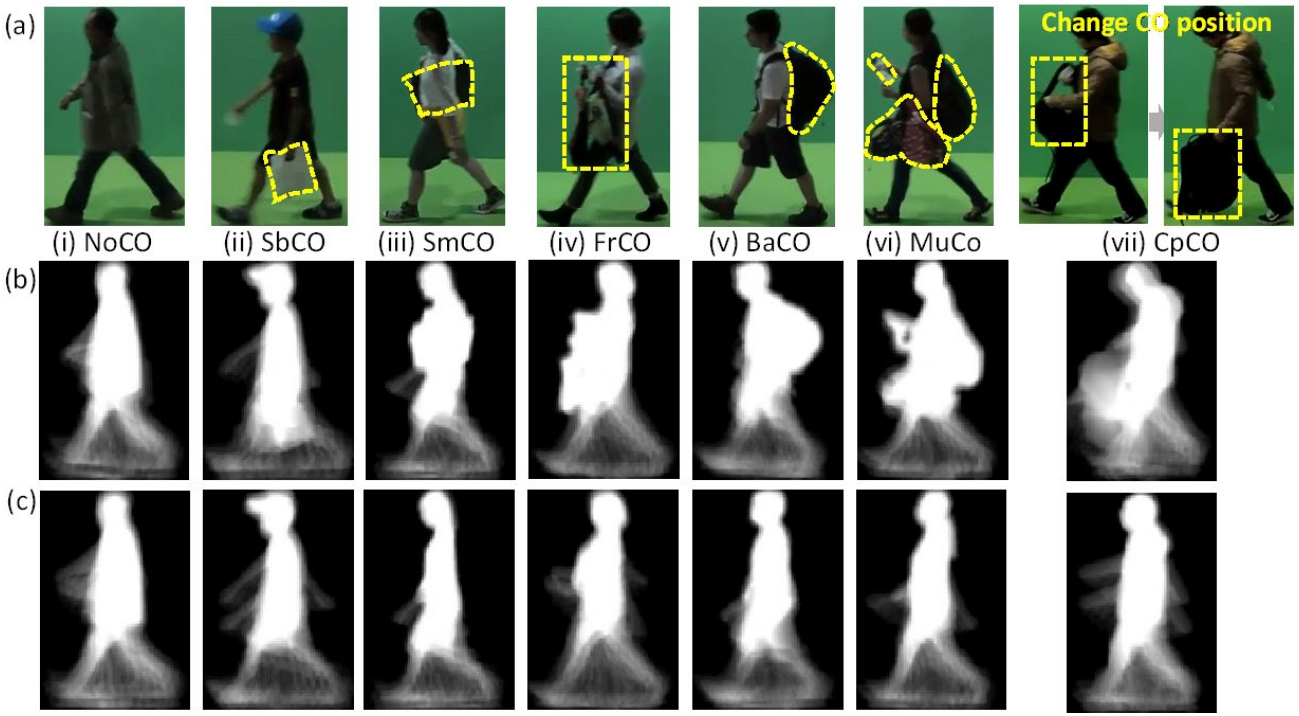
post - estimation

Model-based techniques



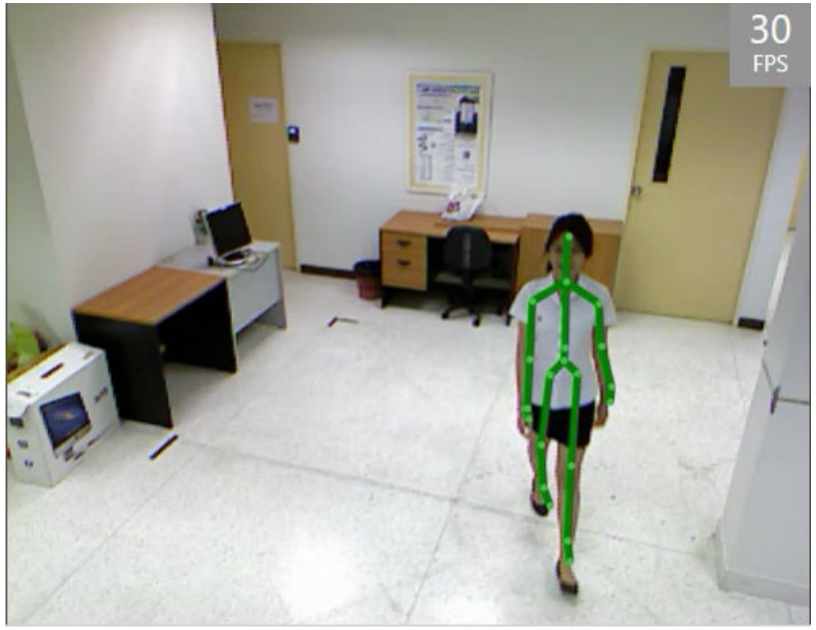
Model-free Gait Recognition

เทคนิคแบบ model-free ใช้ข้อมูลการเดินในรูปแบบ**รูป/วิดีโอ**ที่ได้จากกล้องโดยตรง

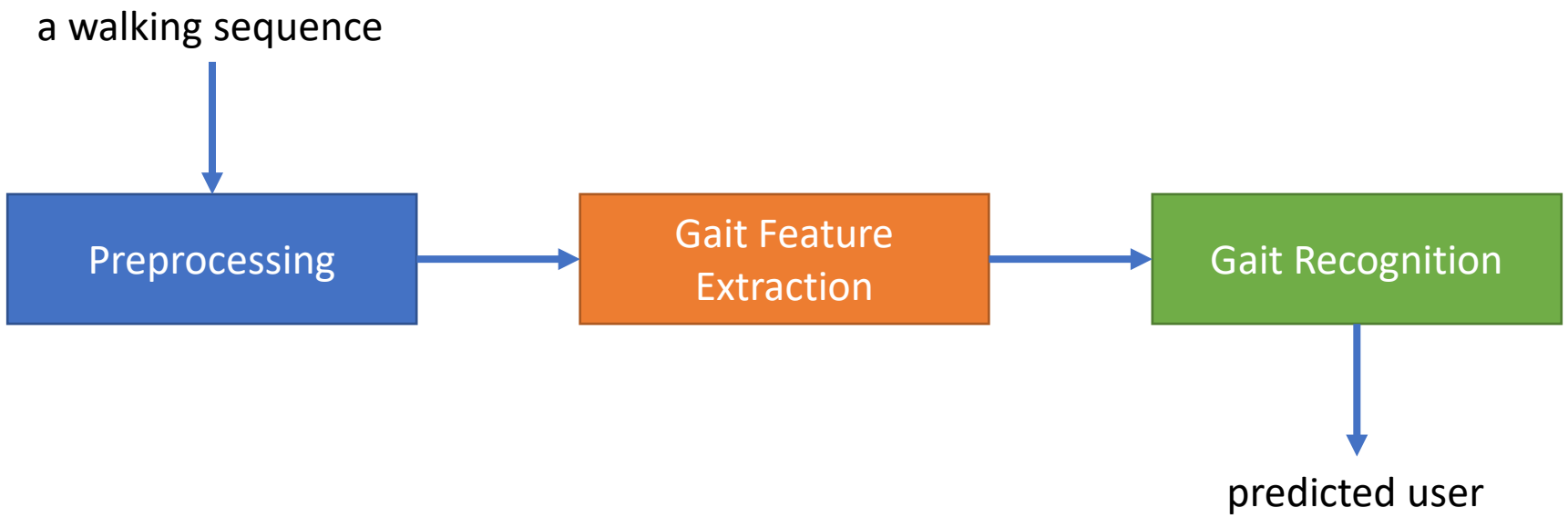


Model-based Gait Recognition

เทคนิคแบบ model-based นำข้อมูลการเดินที่เป็น**รูป/วิดีโอ**มาวิเคราะห์หาโครงสร้างของร่ายกาย แล้วจึงนำข้อมูลโครงสร้างร่างกายมาใช้วิเคราะห์ตัวบุคคล



Gait Recognition Steps



Preprocessing and Gait Feature Extraction

Gait Features

A walking sequence S_i is a sequence of frames:

S_i = Walk

s_i = Frame

Walk

Frame

$$S_i = \langle \mathbf{s}_{i,1}, \mathbf{s}_{i,2}, \ldots, \mathbf{s}_{i,N_i} \rangle \longrightarrow \begin{bmatrix} \quad \end{bmatrix}$$

* Frame · 60

where N_i is the number of frames in the walking sequence S_i. Note that the number of frames varies on different walking sequences. Each frame $\mathbf{s}_{i,j}$ is a vector of 20 joints. Each joint is denoted by a 3D coordinate:

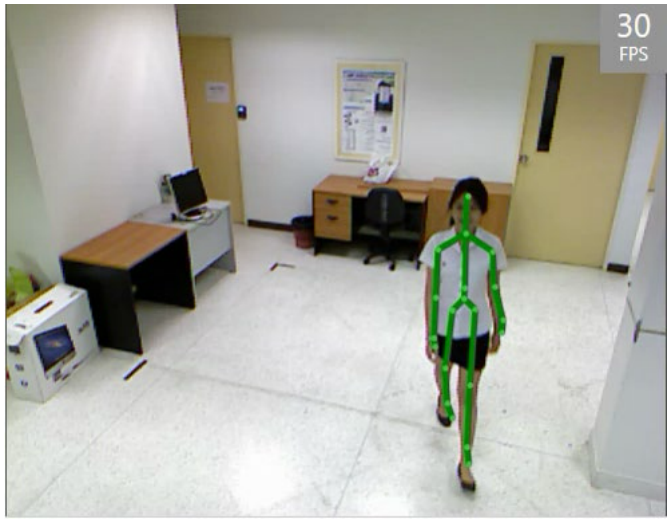
frame

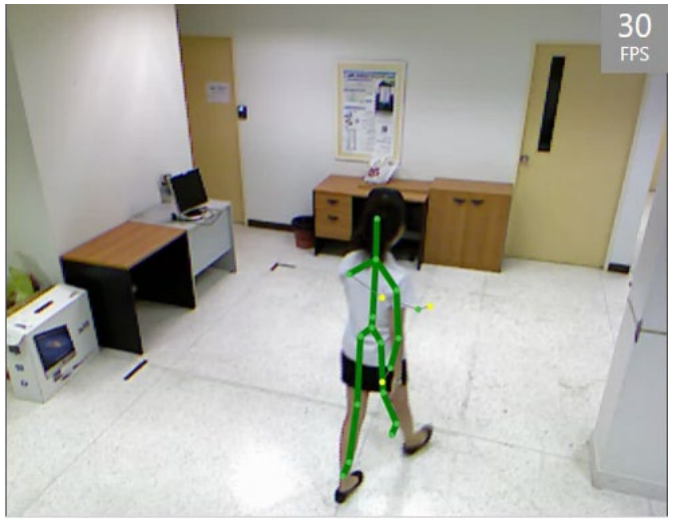
$$\mathbf{s}_{i,j} = \left.\begin{bmatrix} (x_{i,j,1}, y_{i,j,1}, z_{i,j,1}) \\ (x_{i,j,2}, y_{i,j,2}, z_{i,j,2}) \\ \vdots \\ \underbrace{(x_{i,j,20}, y_{i,j,20}, z_{i,j,20})}_{3} \end{bmatrix}\right\} \text{20 joints}$$

* Frame · 20 · 3

Fixed Direction vs Freestyle Walks

- ถ้าบุคคลเดินตรงในทิศทางเดิม และไม่มีการเลี้ยวไปมา จะสามารถนำข้อมูล vector ของแต่ละ frame มาเปรียบเทียบกันได้โดยตรง
- แต่ถ้าบุคคลเดินไม่เป็นเส้นตรง หรือมีการเลี้ยวเปลี่ยนทิศทางในระหว่างการจับภาพ ข้อมูล vector ของแต่ละ frame จะแตกต่างกัน นำมาเปรียบเทียบกับไม่ได้





Center-of-Body Coordinates (1)

ยึดตำแหน่งแกนกลางของร่างกายเป็น Referece

Suppose absolute coordinates (obtained from the Kinect skeletal system) in a particular frame of hip center, spine, right hip and left hip are (x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3) respectively. We define vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ as follow:

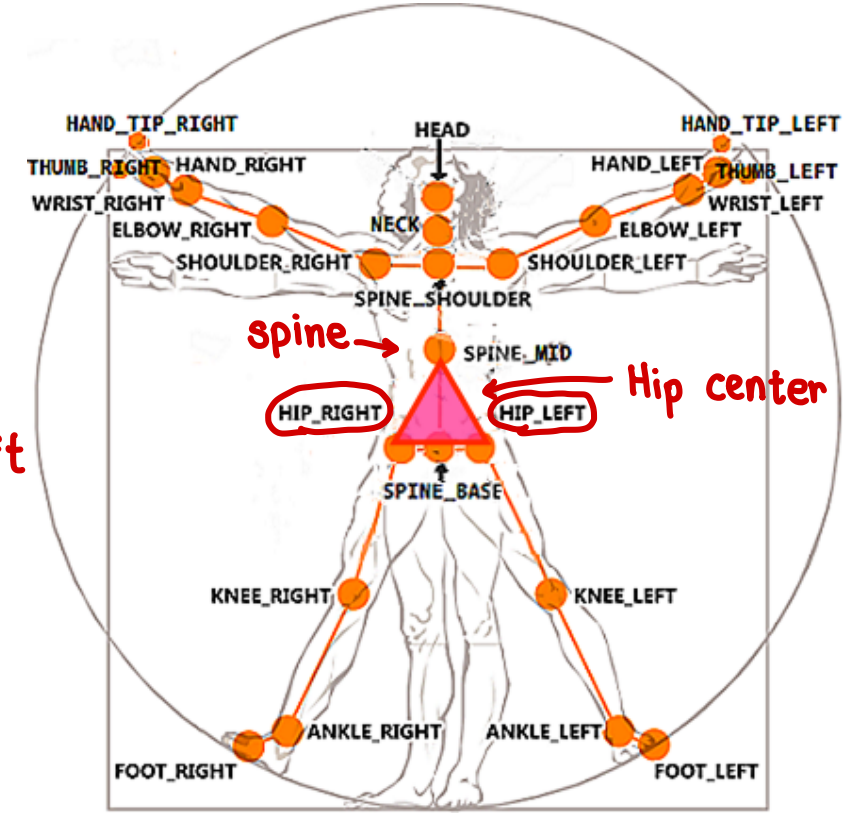
Spine

Hip center

$$\mathbf{v}_1 = \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{bmatrix} \text{ and } \mathbf{v}_2 = \begin{bmatrix} x_3 - x_2 \\ y_3 - y_2 \\ z_3 - z_2 \end{bmatrix}$$

Right → Left

Vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ form a plane that is created from hip center, spine, right hip and left hip with hip center as the origin.



Center-of-Body Coordinates (2)

ถ้า v_1, v_2 ไม่ ⊥ กัน เราจะใช้

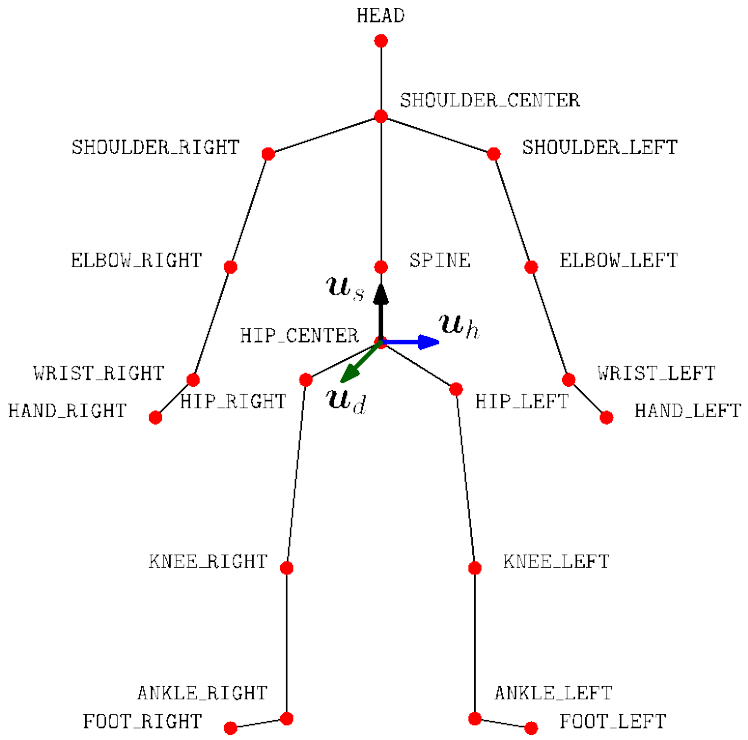
However, the angle between vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ may not be the right angle, we apply Gram-Schmidt process to obtain an orthogonal set of vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that form the same plane. This can be done as follow:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{u}_2 &= \mathbf{v}_2 - \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}_2}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 \end{aligned}$$

A third axis $\mathbf{u}_3$ is defined as

$$\mathbf{u}_3 = \mathbf{u}_2 \times \mathbf{u}_1$$

This third axis is needed to ensure that all the frames are facing in the same direction. The set of vectors $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is a new desired set of axes.



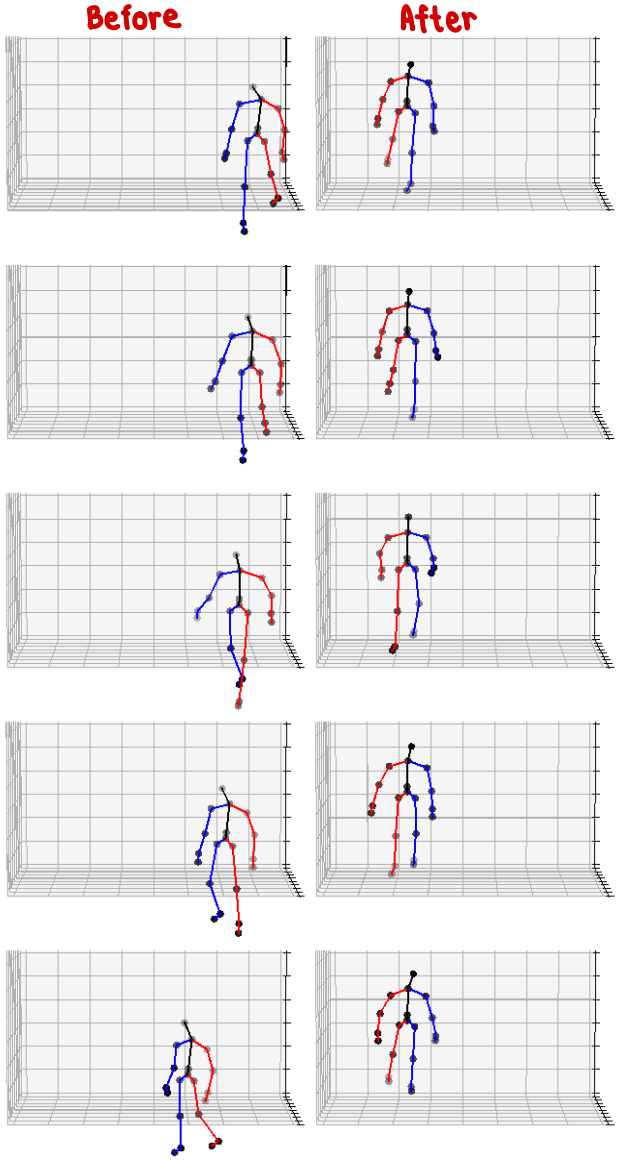
Center-of-Body Coordinates (3)

Suppose an absolute coordinate of a joint in a particular frame is (x, y, z). A relative coordinate of this joint with respect to $\mathcal{B}$ is $(c_1, c_2, c_3)_{\mathcal{B}}$ where c_1, c_2 and c_3 can be calculated from the following equation:

joint [9] - Hip center

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + c_3\mathbf{u}_3$$

These relative coordinates of joints, $(c_1, c_2, c_3)_{\mathcal{B}}$, are then used in the classification process.



Joint Replacement Coordinates

Joint Replacement Coordinates (JRCs) focuses on structures and how three connected joints move in comparison to a middle joint.

joint 3 อันติดกัน = JRCs

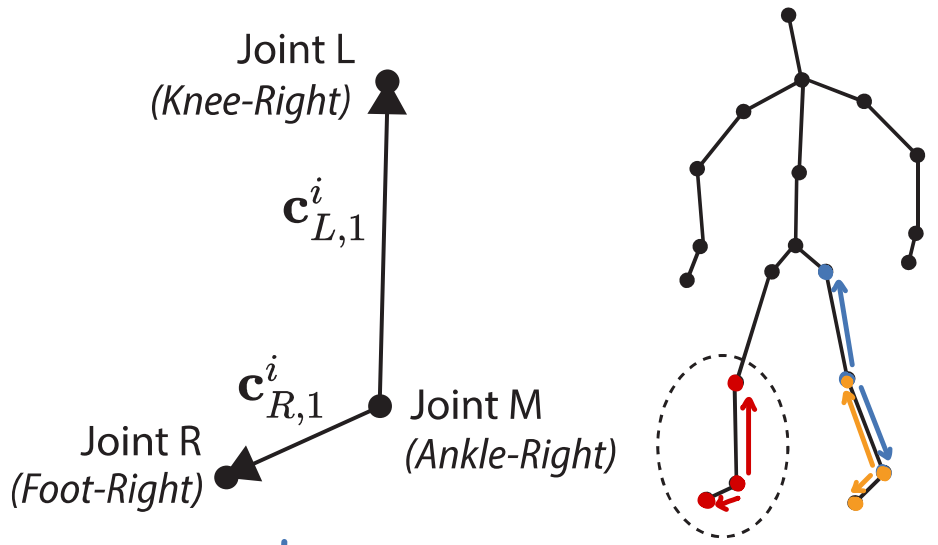
ทั้งหมดมี 22 JRCs

We use 22 JRCs to represent the entire body structure.

JRCs 1 อันจะใช้ vector 2 อันในการ represent

นำ จุดปลาย 1 ลบจุดตรงกลาง = $c^i_{R,1}$

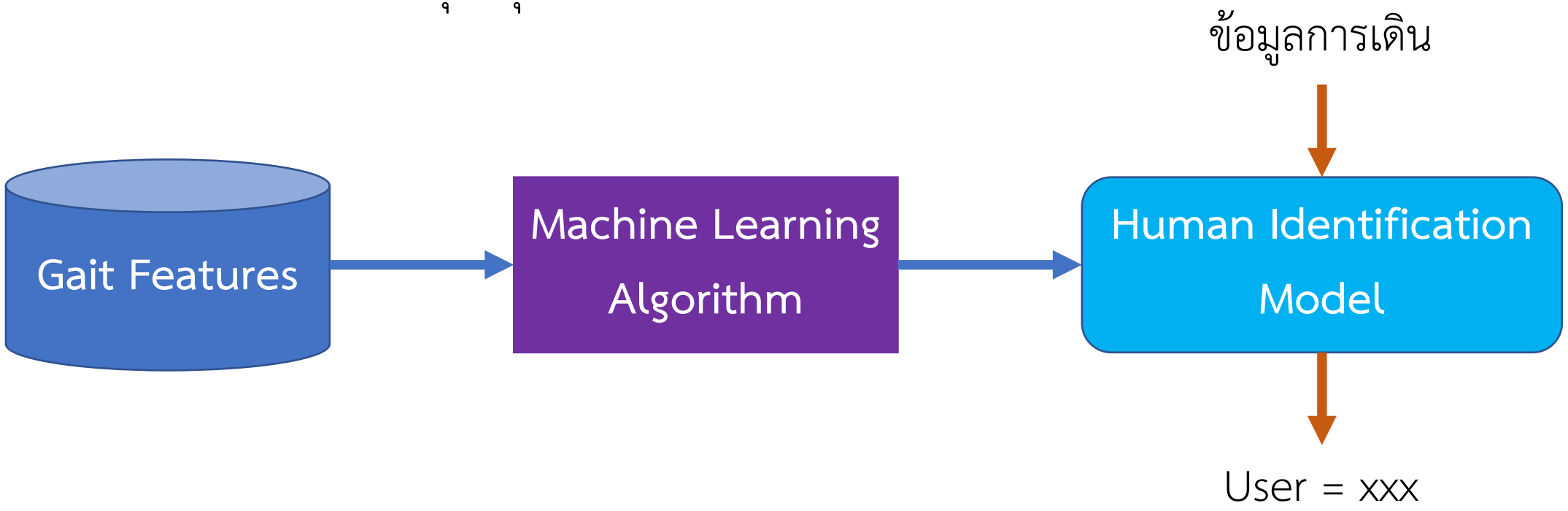
นำ จุดปลาย 2 ลบจุดตรงกลาง = $c^i_{L,1}$



Human Identification Model using Machine Learning Techniques

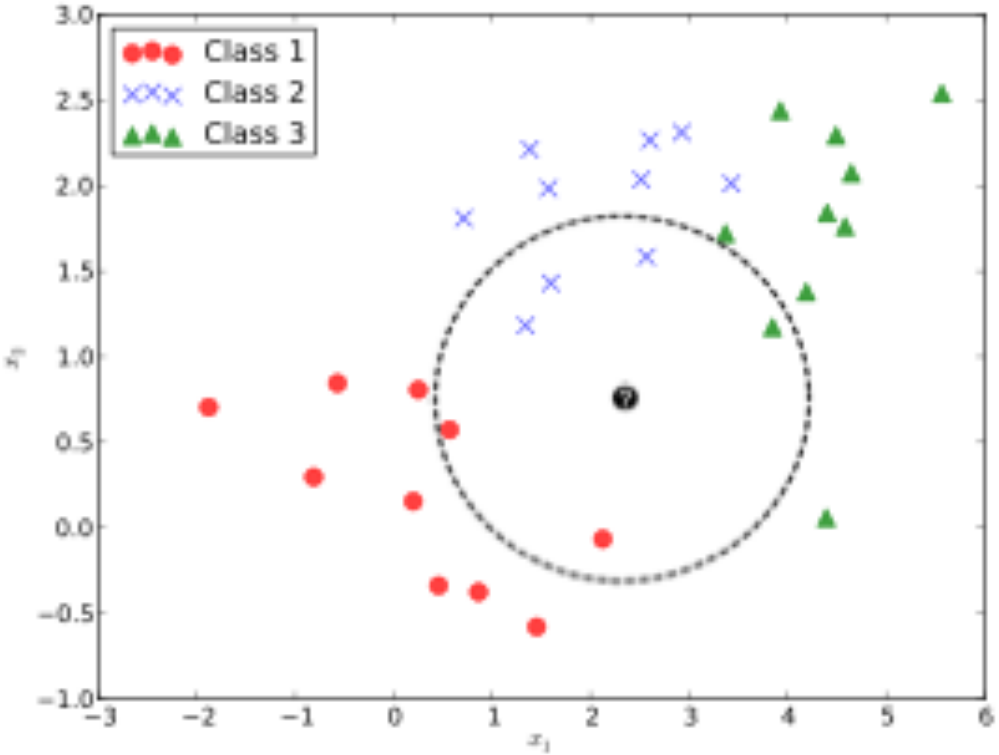
Human Identification Model

คุณลักษณะเฉพาะที่สร้างขึ้นสามารถนำไปใช้ร่วมกับอัลกอริทึมการเรียนรู้ของเครื่อง
เพื่อสร้างโมเดลสำหรับระบุตัวบุคคล



Machine Learning Algorithms

อัลกอริทึมสำหรับการเรียนรู้ของเครื่องมีหลายรูปแบบ เรานำอัลกอริทึมแบบ Supervised Machine Learning หรือการเรียนรู้แบบมีตัวอย่างสอน มาใช้ในกระบวนการ Gait Recognition



Supervised Machine Learning

Given a set of input-output pairs

= matrix # frame 40 x 60 → 20 x 3

Walk

Label : 0 = male
1 = female

$$\mathcal{D} = \{(\mathbf{x}_i, t_i)\}_{i=1}^{N}$$

where,

$\mathcal{D}$ is a training set,

N is the number of examples,

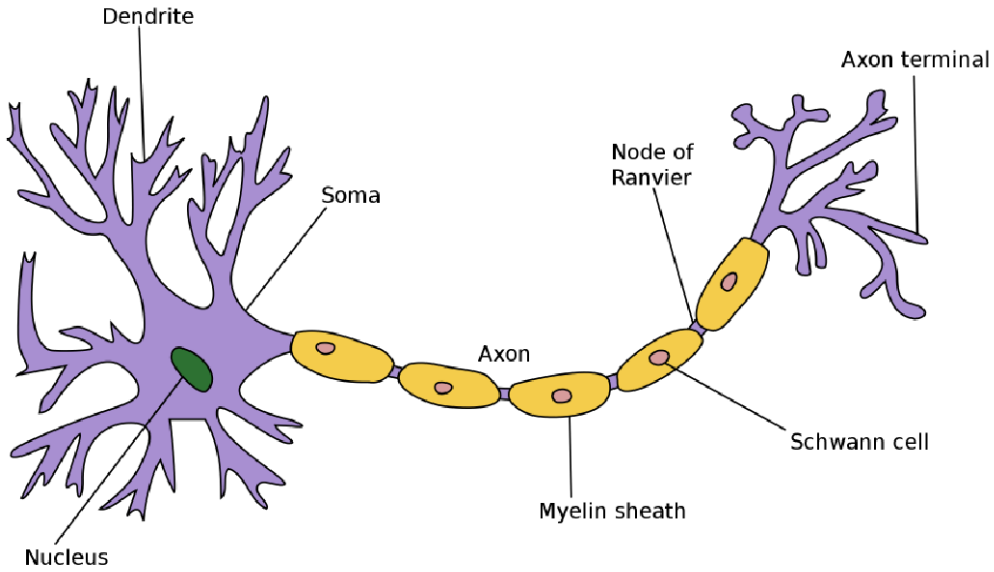
$\mathbf{x}_i \in \mathcal{X}$ is an input vector, and

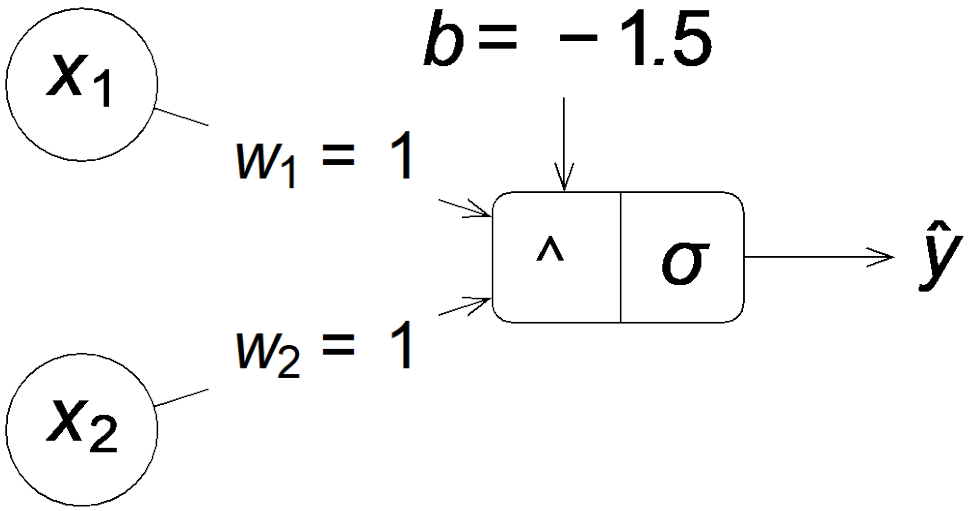
$t_i \in \mathcal{T}$ is a target output,

find a function $\hat{f} : \mathcal{X} \to \mathcal{T}$ that predicts the value of y from an unknown $\mathbf{x}$.

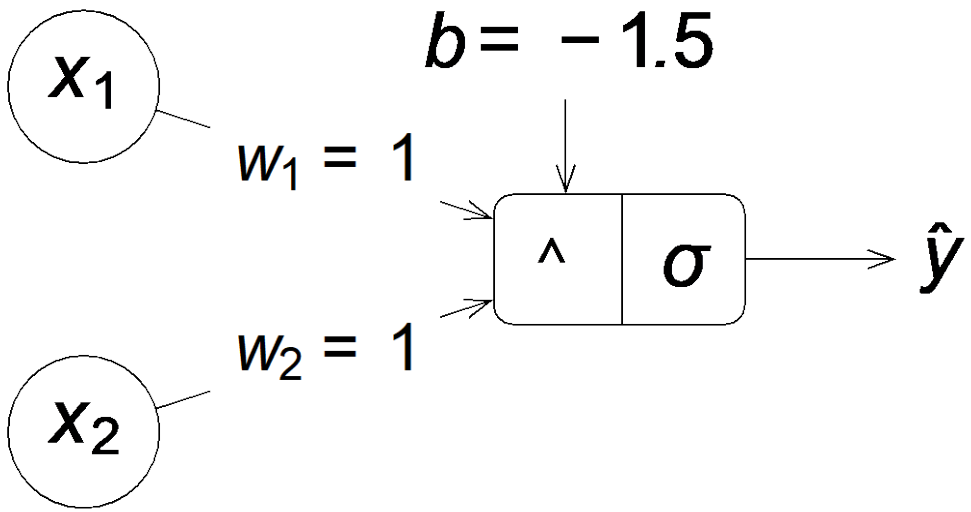
Artificial Neural Network (1)

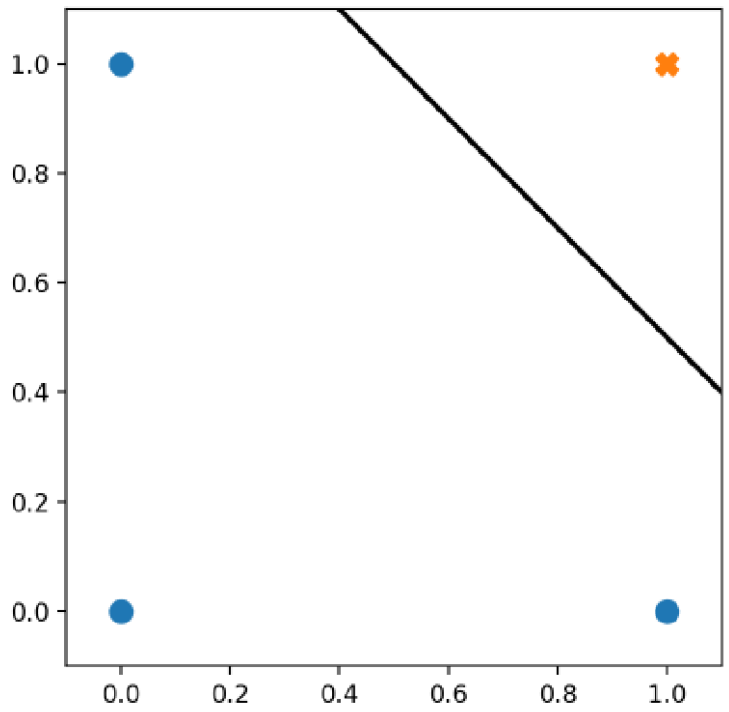
Artificatial Neural Network is a technique inspired by our neurons.





Artificial Neural Network (2)



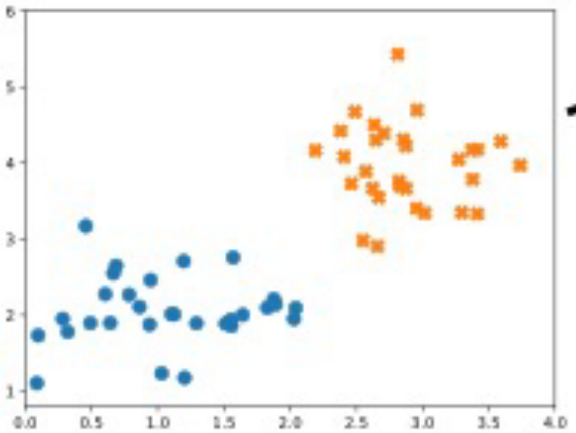


Artificial Neural Network (3)

Structure of

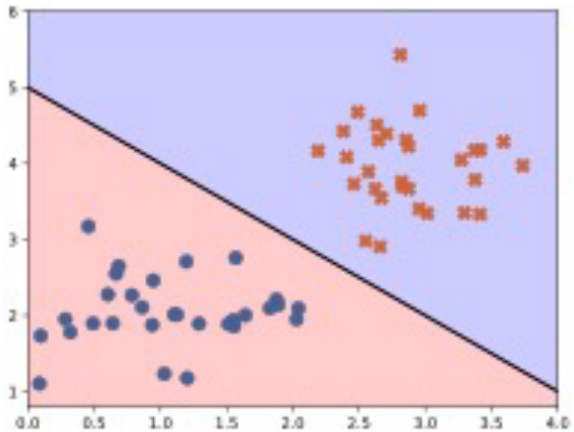
Perceptron

$|\mathbf{w}| = 2$

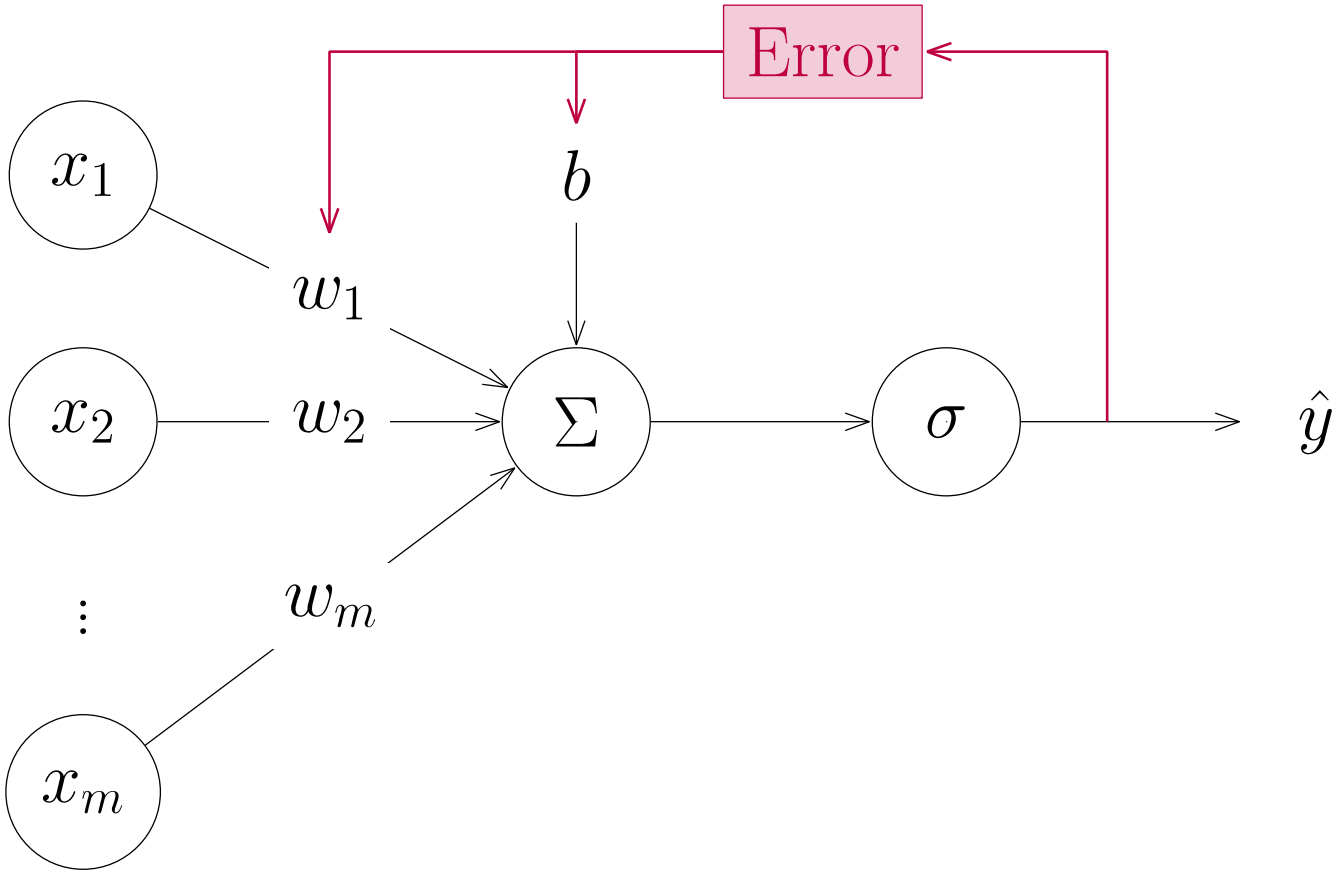


Perceptron Learning Algorithm

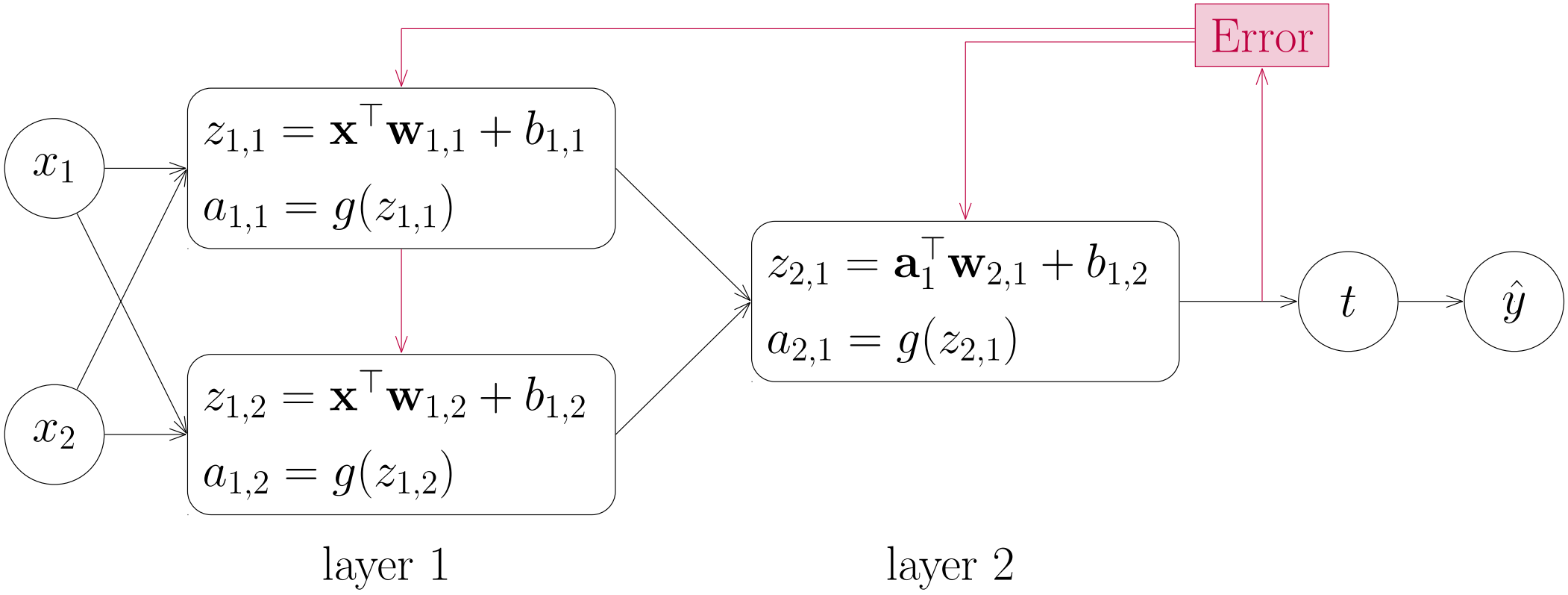
$\mathbf{w}, b$



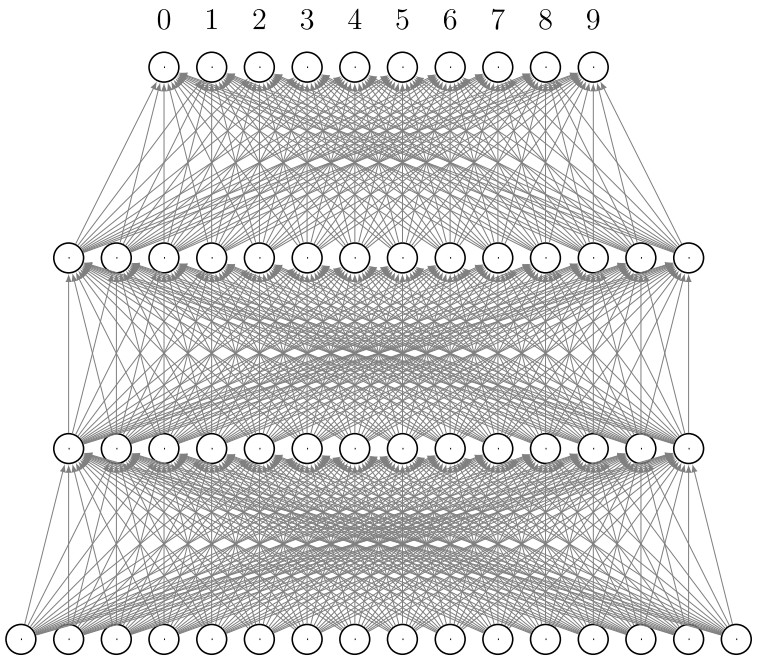
Artificial Neural Network (4)

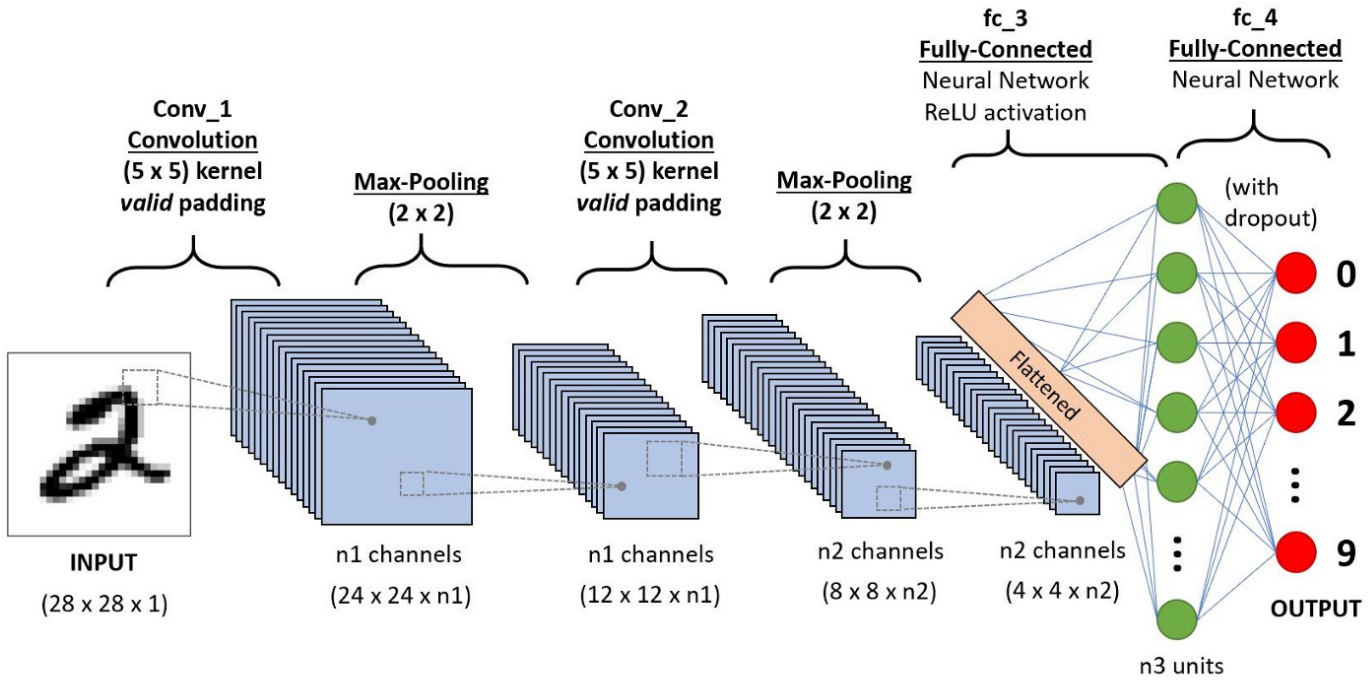


Artificial Neural Network (5)

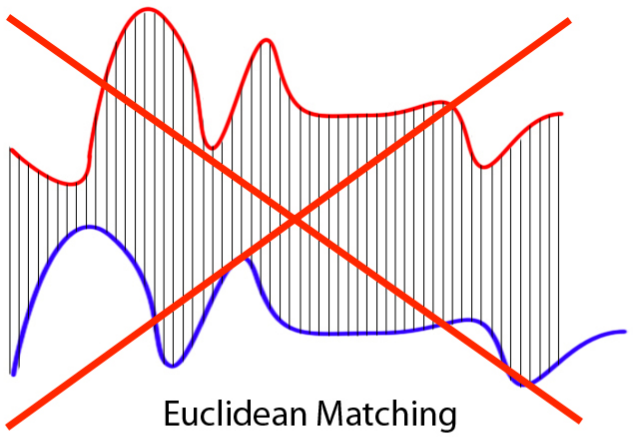


Deep Neural Networks

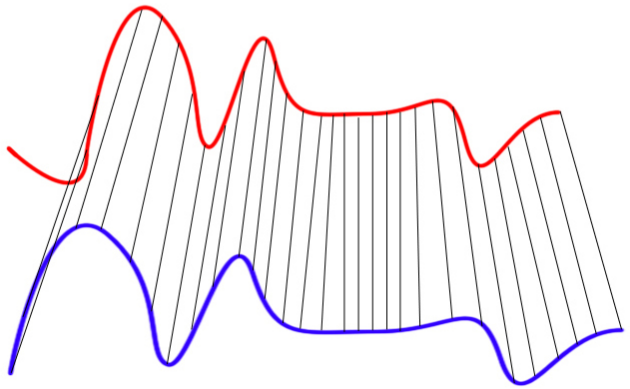




Gait Recognition using Machine Learning



Euclidean Matching



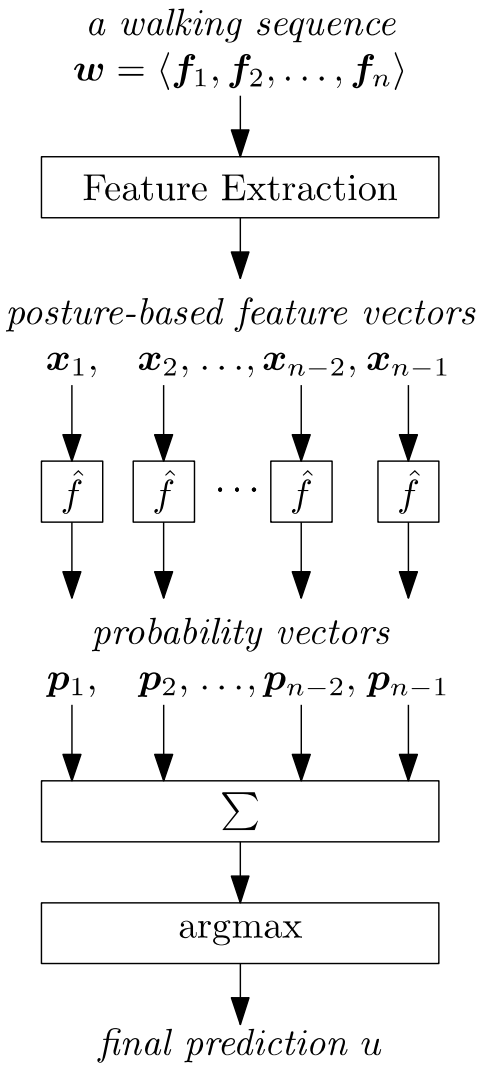
Dynamic Time Warping Matching

- Input ของปัญหานี้ คือ ข้อมูลการเดิน > ลำดับของเฟรมของท่าทาง > ลำดับของ joints
- เนื่องจากลักษณะของการเก็บข้อมูล Input แต่ละตัวจึงยาวไม่เท่ากัน
- การนำการเดิน 2 ครั้งมาเทียบกันตรง ๆ จึงไม่เหมาะสม

Posture-based Gait Recognition

จากการทดลอง เราพบว่าการเดินของบุคคลจะมีจุดเด่นอยู่ที่ท่าทางการเดินในช่วงสั้น ๆ มีรูปร่างเฉพาะของแต่ละบุคคล เราจึงออกแบบเทคนิคที่พิจารณาข้อมูลจากท่าทางในช่วงสั้นๆ แล้วนำผลการทำนายมาผสมรวมกัน เป็นผลการทำนายสุดท้ายของทั้งการเรียน

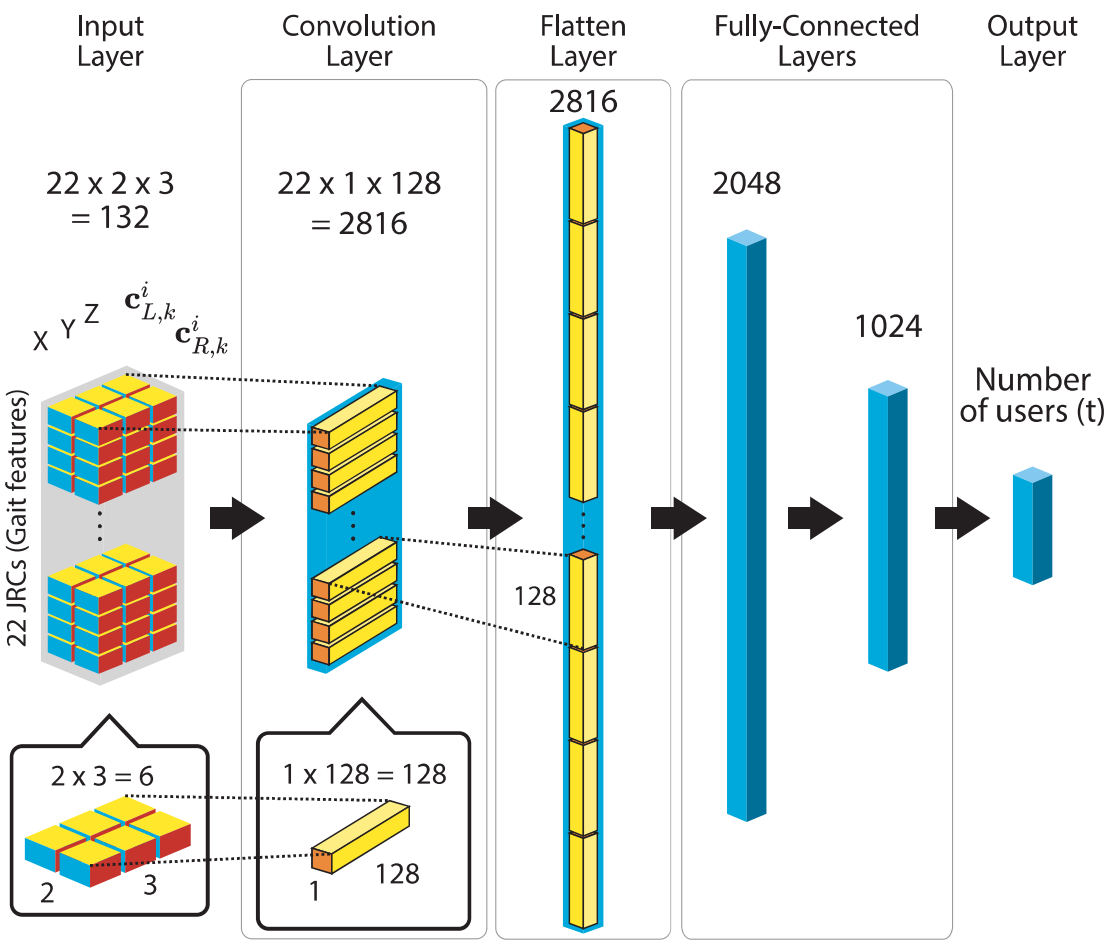
Predice แยก Frame แล้วนามารวมกันโหวตหาค่า predict อีกที



Deep Neural Network based Gait Recognition

เมื่อใช้ JRC เป็น gait feature เราสามารถใช้ Convolutional Neural Network เพื่อสร้างโมเดลสำหรับทำนายบุคคลจากข้อมูลการเดิน

แนวคิดของโครงข่ายนี้ คือ การสร้าง feature เพิ่มเติมจาก JRC แต่ละตัว



ผลการทดลอง (1)

ผลการทดลอง (2)

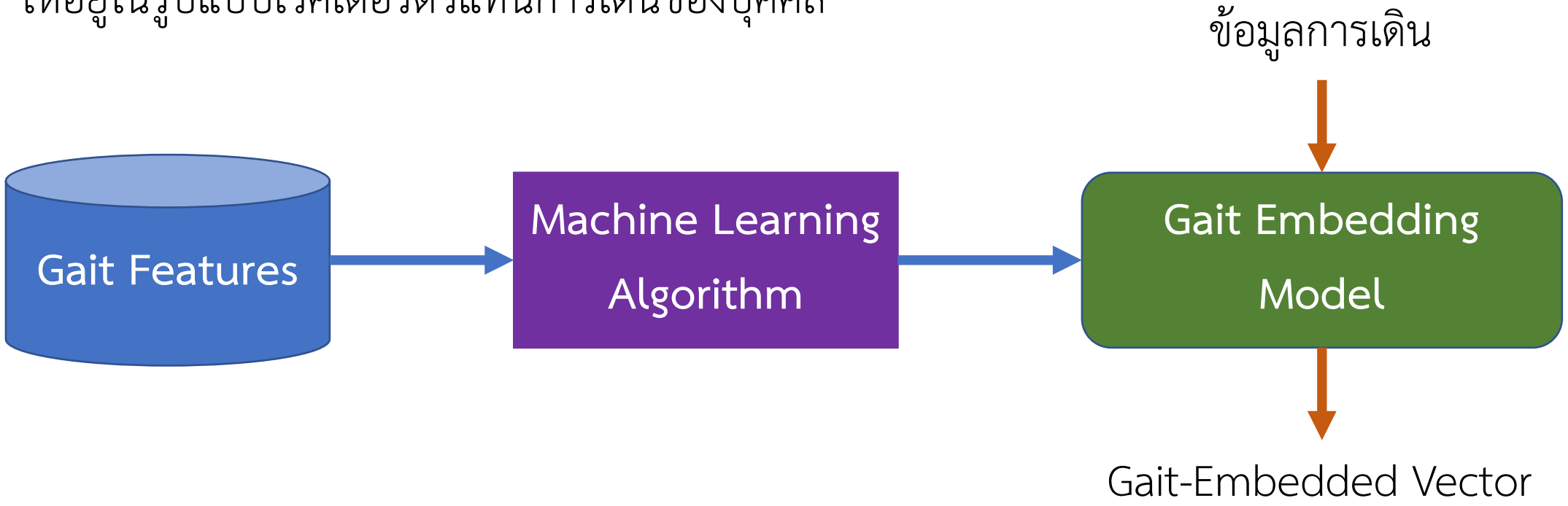
PREDICTION ACCURACIES COMPARING THE PROPOSED TECHNIQUES TO THE EXISTING TECHNIQUES

Techniques	Dataset A 90 Users Free-direction Accuracy (%)	Dataset B 393 Users Free-direction Accuracy (%)	Dataset C 30 Users Fixed-direction Accuracy (%)	Dataset D 483 Users Free-direction and Multi-camera Accuracy (%)	Dataset E 150 Users Free-direction and Multi-camera Accuracy (%)	Dataset F 90 Users Free-direction and Multi-camera Accuracy (%)	Dataset G 50 Users Free-direction and Multi-camera Accuracy (%)
JRC-CNN	$\mathbf{98.4 \pm 1.3}$	$\mathbf{98.4 \pm 0.6}$	99.3 ± 1.5	$\mathbf{98.4 \pm 0.6}$	$\mathbf{98.9 \pm 0.8}$	$\mathbf{98.4 \pm 1.0}$	$\mathbf{99.4 \pm 0.5}$
JRC-KNN	97.4 ± 1.4	$95.6 \pm 0.5^{\dagger}$	98.0 ± 3.0	$95.8 \pm 1.0^{\dagger}$	$97.7 \pm 0.8^{\dagger}$	98.2 ± 1.7	99.3 ± 0.6
JRC-MLP	$96.8 \pm 1.2^{\ddagger}$	$92.6 \pm 1.3^{\dagger}$	99.3 ± 1.5	$93.2 \pm 0.8^{\dagger}$	$96.2 \pm 1.5^{\dagger}$	$96.2 \pm 1.7^{\ddagger}$	$98.1 \pm 1.4^{\ddagger}$
CoB-CNN [28]	$97.3 \pm 1.3^{\dagger}$	$93.0 \pm 1.2^{\dagger}$	98.7 ± 3.0	$93.4 \pm 0.9^{\dagger}$	$96.9 \pm 1.3^{\dagger}$	$96.0 \pm 1.9^{\ddagger}$	$98.3 \pm 1.3^{\ddagger}$
CoB-KNN [28]	$96.2 \pm 1.4^{\dagger}$	$93.6 \pm 0.7^{\dagger}$	98.7 ± 3.0	$93.4 \pm 1.0^{\dagger}$	$95.5 \pm 1.0^{\dagger}$	$94.7 \pm 1.4^{\dagger}$	$97.2 \pm 1.3^{\dagger}$
CoB-MLP [28]	$97.4 \pm 1.2^{\dagger}$	$92.4 \pm 1.7^{\dagger}$	98.7 ± 1.8	$93.6 \pm 0.9^{\dagger}$	$95.7 \pm 1.0^{\dagger}$	$91.3 \pm 2.0^{\dagger}$	$85.0 \pm 1.8^{\dagger}$
Choi et al. [29] [a]	$91.9 \pm 2.5^{\dagger}$		$\mathbf{100.0 \pm 0.0}$		$91.9 \pm 2.5^{\dagger}$	$94.0 \pm 1.9^{\dagger}$	
Yang et al. [19]	$55.1 \pm 3.8^{\dagger}$	$43.9 \pm 1.6^{\dagger}$	$94.0 \pm 3.7^{\ddagger}$	$43.9 \pm 1.3^{\dagger}$	$57.8 \pm 2.5^{\dagger}$	$68.2 \pm 6.6^{\dagger}$	$82.4 \pm 5.0^{\dagger}$
Han et al. [6] [b]	$80.4 \pm 3.4^{\dagger}$	$71.2 \pm 0.8^{\dagger}$		$56.3 \pm 1.4^{\dagger}$	$76.9 \pm 2.3^{\dagger}$		$98.2 \pm 1.0^{\dagger}$

Gait Embedding Model

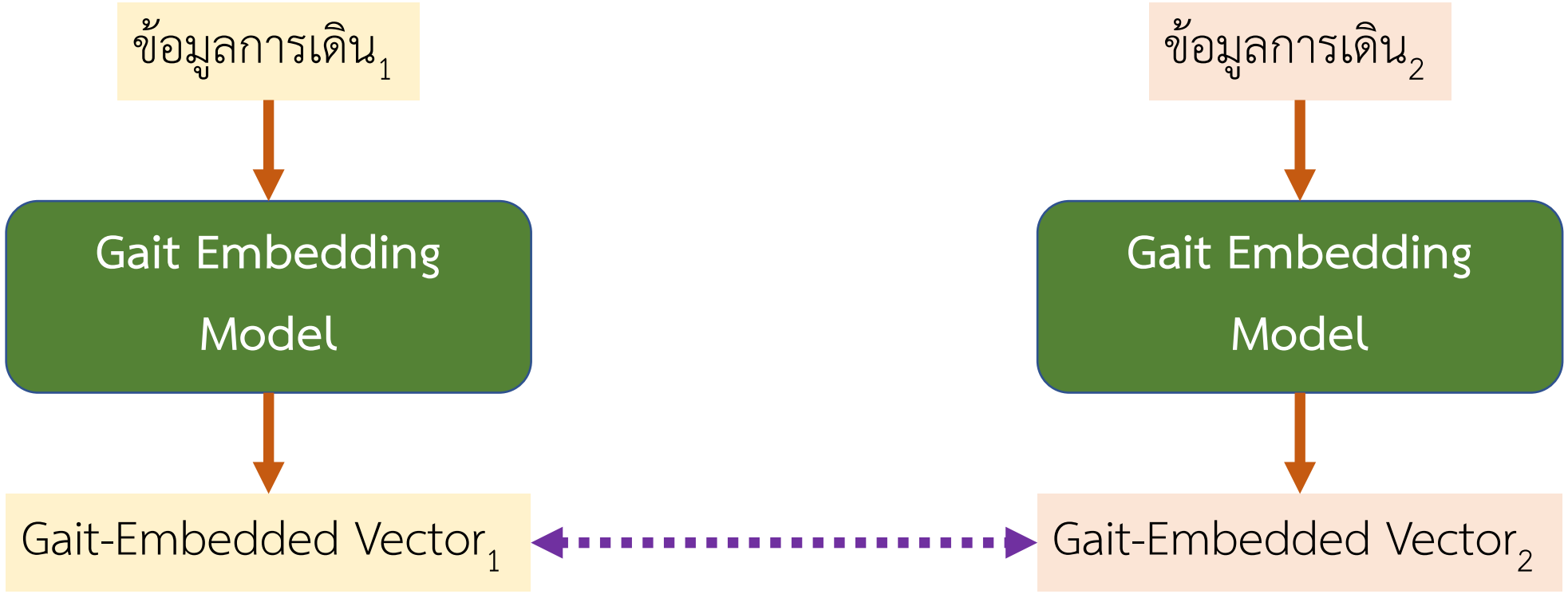
Gait Embedding Model

คุณลักษณะเฉพาะสามารถนำไปใช้สร้างโมเดลเพื่อแปลงท่าทางการเดินของบุคคล
ให้อยู่ในรูปแบบเวคเตอร์ตัวแทนการเดินของบุคคล



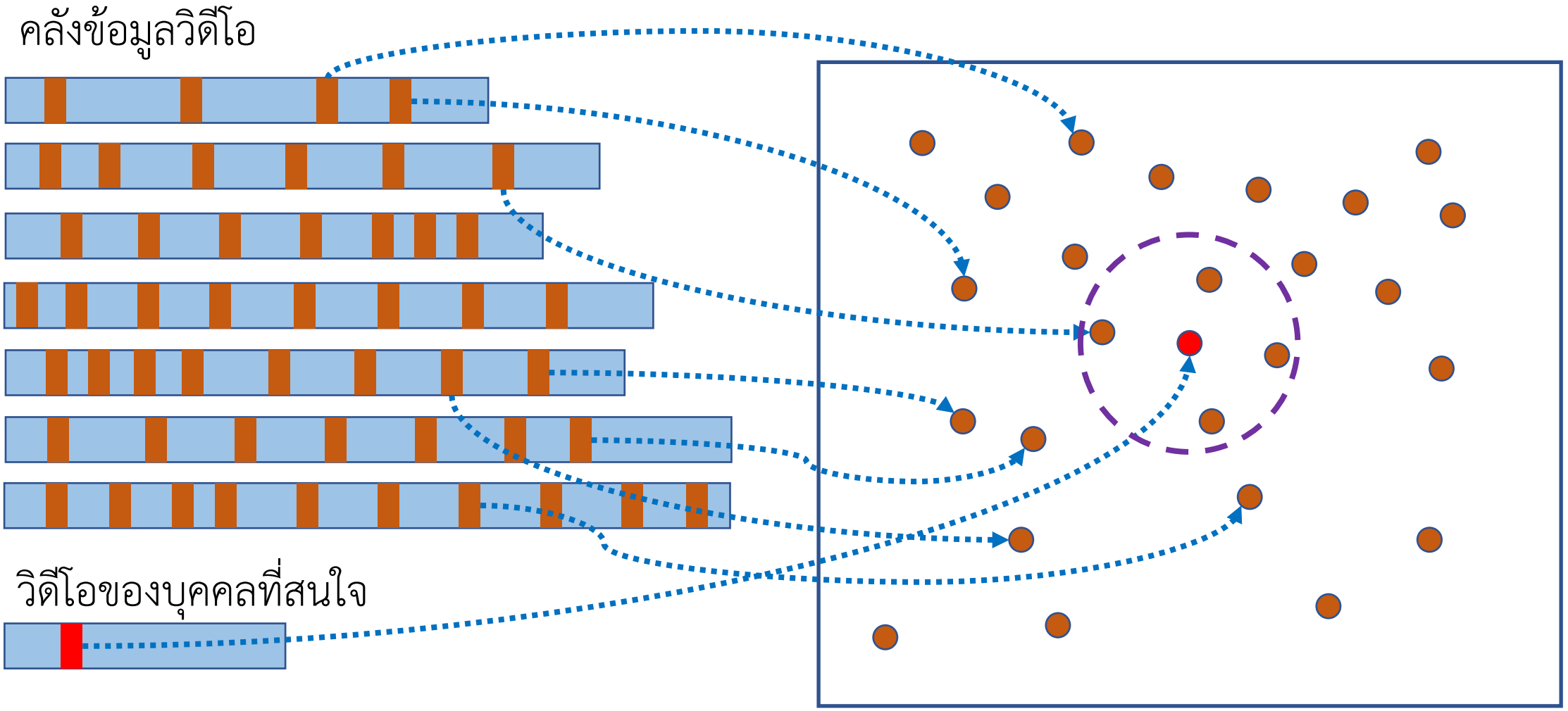
Gait Embedded Vectors

Gait-Embedded Vector เป็นเหมือนลายเซ็นแทนท่าทางการเดินของแต่ละคน

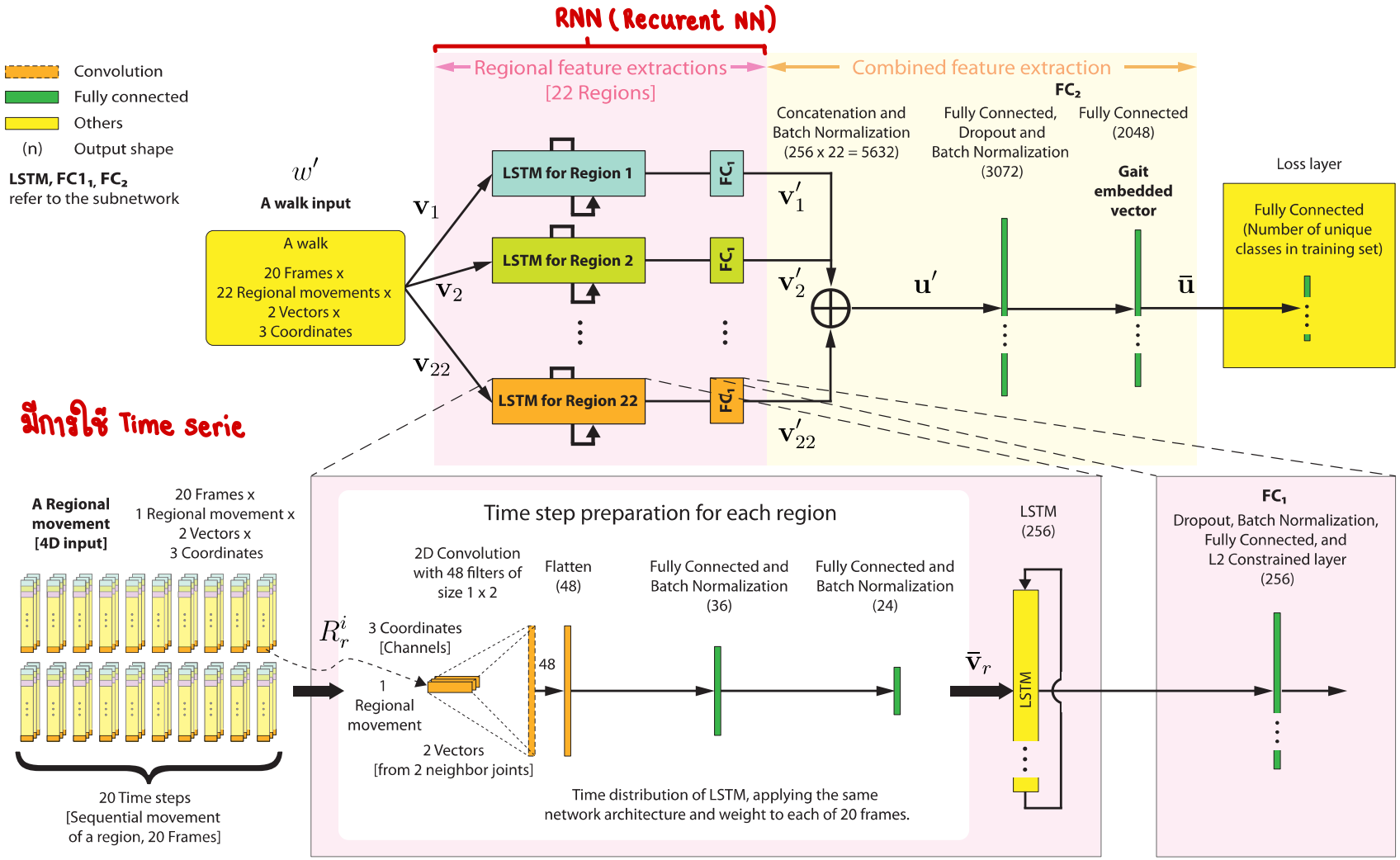


ความคล้ายของท่าทางการเดิน

Search with Gait Embedded Vectors



Deep Neural Network for Gait Embedding



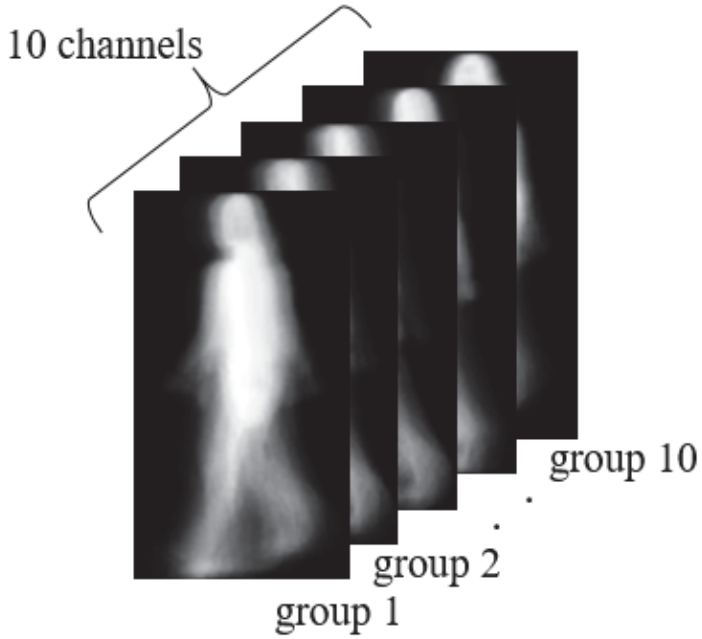
ผลการทดลอง (3)

Techniques	Top-k Accuracy (%) 1	2	3	4	5	Mean AUC (%) ROC	PR
Regional LSTM with L2 and Softmax	**77.4 ± 1.6**	**85.3 ± 1.2**	**88.8 ± 0.9**	**90.8 ± 0.7**	**92.2 ± 0.7**	86.9 ± 0.9	19.6 ± 1.5
Regional LSTM with L2 and AdaCos	57.7† ± 3.5	69.6† ± 2.8	75.5† ± 2.4	79.3† ± 2.0	81.9† ± 1.8	87.6 ± 0.9	18.8 ± 1.4
Regional LSTM with Softmax	76.6 ± 1.7	84.7 ± 1.3	88.2 ± 1.0	90.4 ± 0.9	91.9 ± 0.7	87.4 ± 1.0	**20.1 ± 1.5**
Regional LSTM with AdaCos	56.2† ± 3.1	68.5† ± 2.4	74.6† ± 2.0	78.4† ± 1.8	81.1† ± 1.6	**87.7 ± 1.0**	19.1 ± 1.6
Regional LSTM with AMSoftmax	65.7† ± 5.4	76.1† ± 4.3	81.1† ± 3.6	84.1† ± 3.1	86.2† ± 2.8	86.7† ± 1.1	18.6 ± 1.5
LSTM with L2 and Softmax	39.8† ± 1.8	53.3† ± 1.9	60.9† ± 1.8	66.0† ± 1.7	69.7† ± 1.6	85.2† ± 0.9	13.9† ± 1.0
LSTM with L2 and AdaCos	33.2† ± 2.0	46.6† ± 2.1	54.6† ± 2.0	60.3† ± 1.9	64.5† ± 1.8	84.5† ± 0.8	12.6† ± 0.8
LSTM with Softmax	32.9† ± 6.7	46.0† ± 8.0	53.9† ± 8.3	59.4† ± 8.3	63.6† ± 8.1	83.9† ± 2.6	12.1† ± 2.2
LSTM with AdaCos	32.5† ± 3.1	46.1† ± 3.3	54.2† ± 3.2	59.8† ± 3.1	64.1† ± 2.9	84.4† ± 1.3	12.5† ± 1.3
Han et al. (GEI) [12]	67.1† ± 1.7	75.2† ± 1.3	79.2† ± 1.2	81.9† ± 1.1	83.8† ± 1.0	61.6† ± 0.7	4.8† ± 0.2
Yang et al. [20]	34.0† ± 2.2	45.3† ± 2.1	52.2† ± 2.0	57.1† ± 2.0	61.0† ± 1.9	63.5† ± 2.1	4.1† ± 0.4
Andersson et al. [19]	29.8† ± 1.9	40.9† ± 2.0	47.8† ± 2.1	52.9† ± 2.0	57.0† ± 2.0	65.3† ± 2.3	4.5† ± 0.5

Other Related Works

Gender Classification from Gait Silhouette

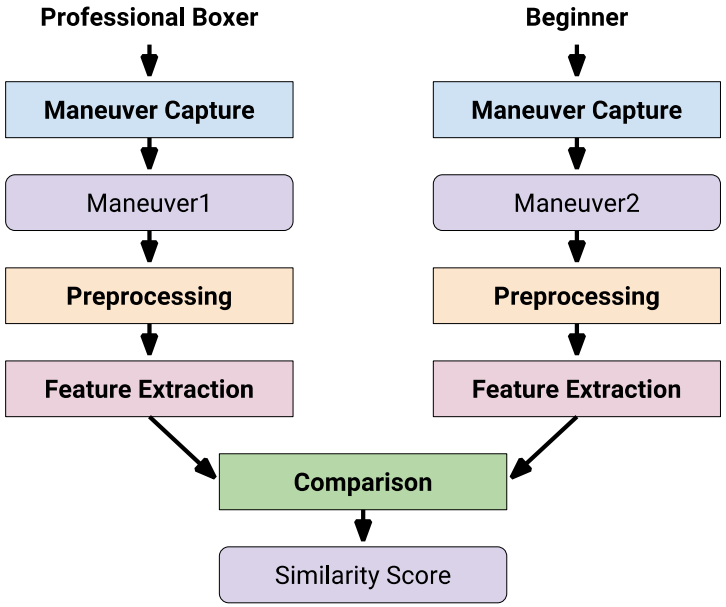
- ทำนายเพศของบุคคลจากภาพที่จับได้ในวิดีโอ
- ทำงานแบบ appearance-based และรองรับการเดินแบบ free-style
- สร้างคุณลักษณะเฉพาะโดยใช้ Machine Learning เทียบทิศทางการเดินในแต่ละช่วงเวลา เพื่อสร้างกลุ่มภาพแทนบุคคล
- ทำนายเพศของบุคคลได้ด้วยความแม่นยำ 97.58%

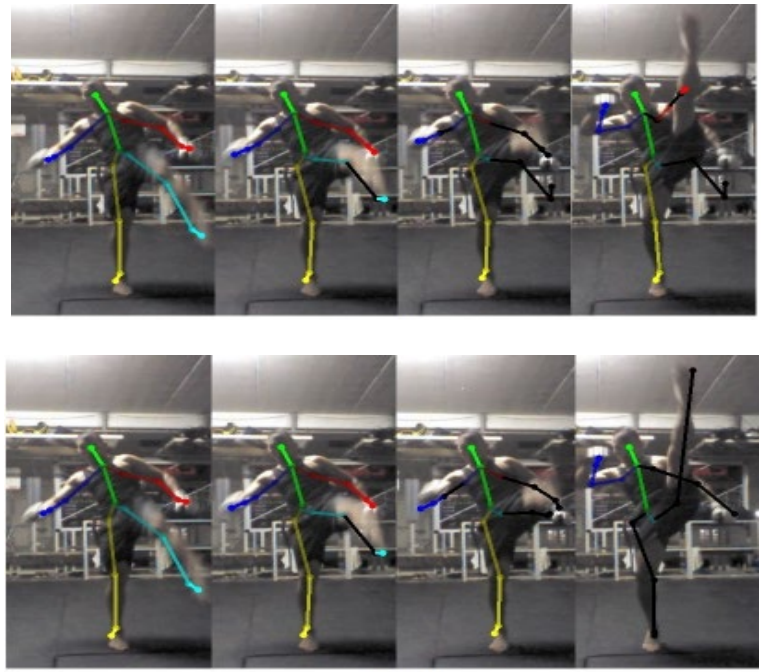


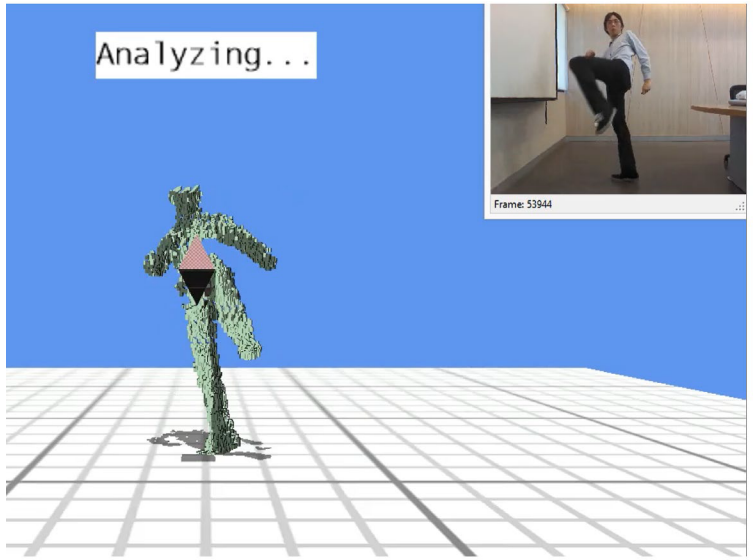
K. Kitchat, N. Khamsemanan, and C. Nattee, “Gender classification from gait silhouette using observation angle-based GEIs,” IEEE Conference on Cybernetics and Intelligence Systems (CIS) and Robotics, Automation and Mechatronics (RAM), 2019.

Instructional Support System for Muay Thai

- ข้อมูลโครงสร้างร่างกายจาก Kinect สามารถนำมาใช้สร้างเกมเพื่อฝึกแม่ไม้มวยไทย โดยเก็บข้อมูลท่าจากนักมวยอาชีพ เพื่อใช้เป็นมาตรฐานสำหรับเปรียบเทียบกับผู้เล่น





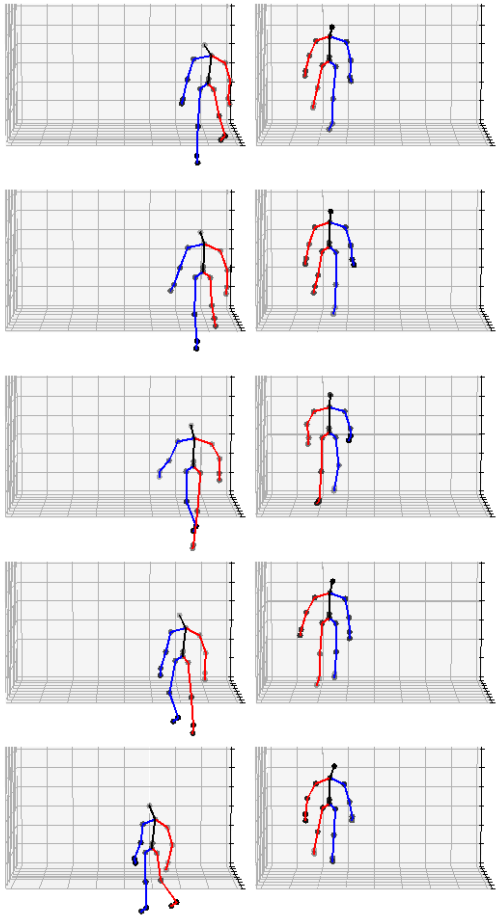


References

- N. Khamsemanan, C. Nattee, and N. Jianwattanapaisarn, “Human identification from freestyle walks using posture-based gait feature,” IEEE Trans. Inf. Forensics Security, vol. 13, no. 1, pp. 119-128, Jan. 2018.
- P. Limcharoen, N. Khamsemanan, and C. Nattee, “View-independent gait recognition using joint replacement coordinates (JRCs) and convolutional neural network,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 3430-3442, 2020.
- Limcharoen, P., Khamsemanan, N., Nattee, C. (2021). Gait Recognition and Re-Identification Based on Regional LSTM for 2-Second Walks. IEEE Access, Volume 9, 2021, Digital Object Identifier 10.1109/ACCESS.2021.3102936.

ตัวอย่างผลงานวิจัยที่เกี่ยวข้องโดยตรง

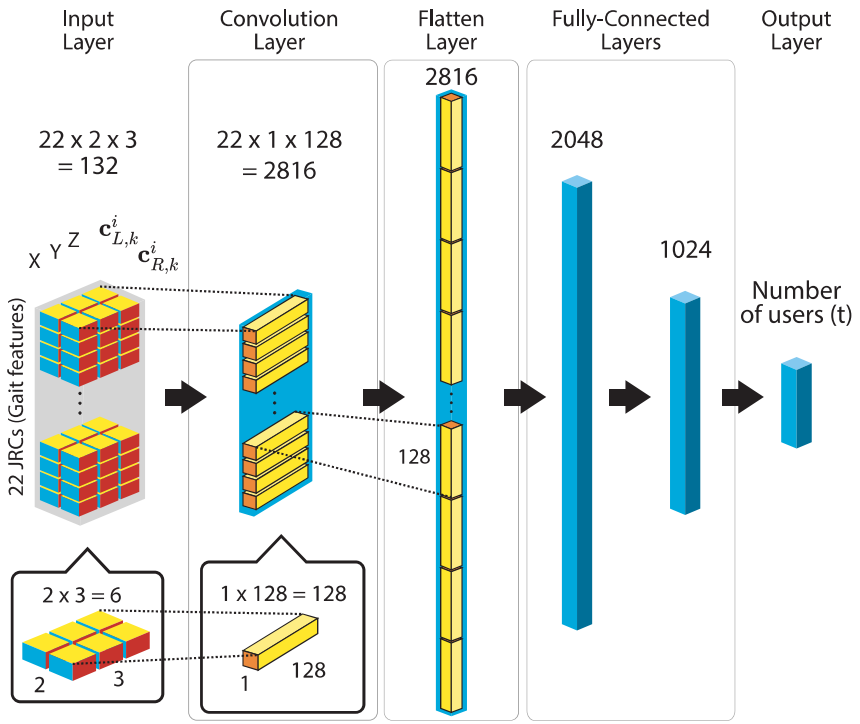
- Posture-based Gait Features
  - เทคนิคสำหรับ gait recognition แบบ model-based
  - ข้อมูลการเดินที่นำมาใช้เป็นแบบ free-style walk ไม่จำเป็นต้องเดินเป็นเส้นตรง หรือเดินตามทางที่กำหนด
  - ใช้การวิเคราะห์เชิงเรขาคณิต เพื่อหมุนโครงร่างของบุคคลให้อยู่ในทิศเดียวกัน
  - ทำนายบุคคลได้ด้วยความแม่นยำ 96.8%



N. Khamsemanan, C. Nattee, and N. Jianwattanapaisarn, “Human identification from freestyle walks using posture-based gait feature,” IEEE Trans. Inf. Forensics Security, vol. 13, no. 1, pp. 119-128, Jan. 2018.

ตัวอย่างผลงานวิจัยที่เกี่ยวข้องโดยตรง

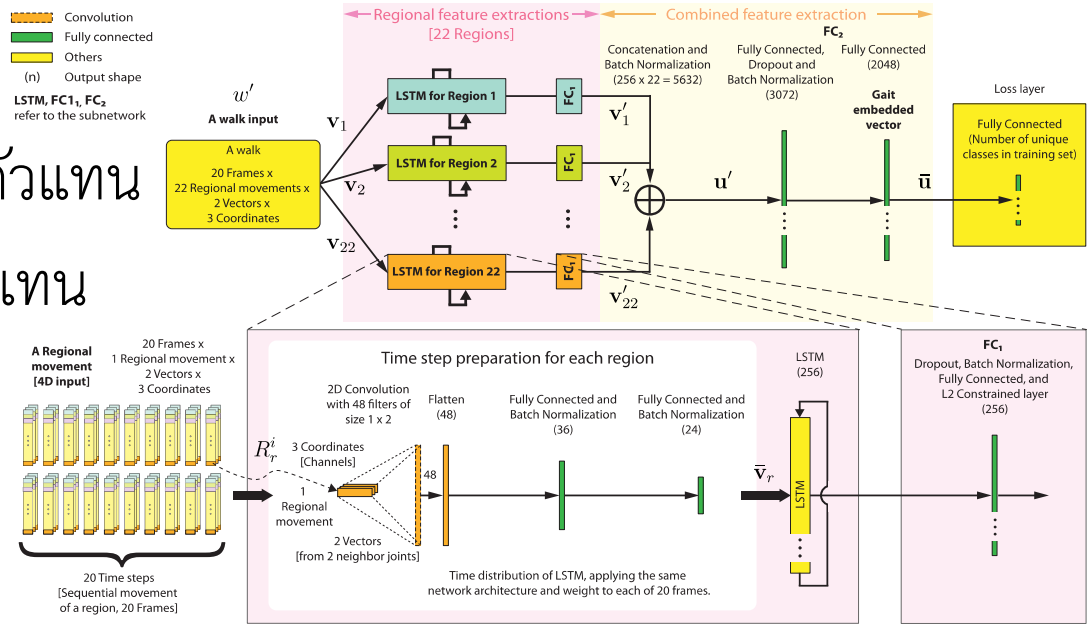
- View-Independent Gait Recognition using Joint Replacement Coordinates (JRCs) and Convolutional Neural Network
  - เสนอวิธีสร้างคุณลักษณะเฉพาะแบบใหม่ ที่สามารถดึงคุณลักษณะจากท่าทางการเดินโดยไม่ขึ้นกับมุมกล้อง
  - ใช้ Deep Learning ในการสร้างโมเดลเพื่อทำนายบุคคล
  - ทำนายบุคคลจำนวนมาก ได้ด้วยความแม่นยำ 99.4%



P. Limcharoen, N. Khamsemanan, and C. Nattee, “View-independent gait recognition using joint replacement coordinates (JRCs) and convolutional neural network,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 3430-3442, 2020.

ตัวอย่างผลงานวิจัยที่เกี่ยวข้องโดยตรง

- Gait Recognition and Re-Identification
  - นำเสนอ Gait Embedding Model เพื่อแปลงข้อมูลท่าทางการเดิน 2 วินาทีให้อยู่ในรูปเวคเตอร์ตัวแทน
  - สามารถหาใช้วิธีการหาความคล้ายของเวคเตอร์ตัวแทนเพื่อเปรียบเทียบว่าการเดิน 2 ครั้งเป็นของบุคคลเดียวกันหรือไม่
  - ดึงข้อมูลบุคคลเดียวกันจากคลังข้อมูลได้ด้วยความแม่นยำ 92.2%



Limcharoen, P., Khamsemanan, N., Nattee, C. (2021). Gait Recognition and Re-Identification Based on Regional LSTM for 2-Second Walks. IEEE Access, Volume 9, 2021, Digital Object Identifier 10.1109/ACCESS.2021.3102936.