

## Phương pháp dựa trên học máy trong theo dõi chuyển động cơ thể người với số lượng ít cảm biến

Đặng Hoàng Minh\*, Phùng Như Hải, Lưu Văn Sáng, Vũ Hoàng Minh

Viện Công nghệ thông tin/Viện KH-CN quân sự.

\*Email: danghoangminh86@gmail.com

Nhận bài: 30/8/2022; Hoàn thiện: 8/11/2022; Chấp nhận đăng: 28/11/2022; Xuất bản: 20/12/2022.

DOI: <https://doi.org/10.54939/1859-1043.j.mst.FEE.2022.171-176>

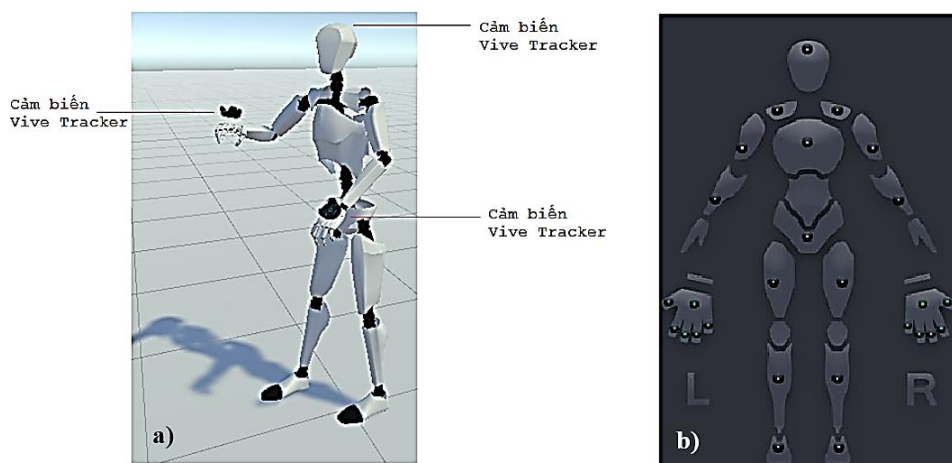
### TÓM TẮT

Hầu hết các bộ thiết bị cảm biến cơ thể hiện nay được cấu tạo bằng nhiều cảm biến con quay kết hợp cảm biến gia tốc (IMU) đặt trên các vị trí khác nhau của cơ thể người. Các cảm biến IMU sau đó sẽ cảm biến vị trí, góc quay, góc nghiêng của mình trong không gian, từ đó, nội suy ra chuyển động của các bộ phận và toàn bộ cơ thể người. Mặc dù cảm biến IMU có độ chính xác và tốc độ xử lý nhanh, các cảm biến loại này chịu một hạn chế lớn là dễ bị ảnh hưởng bởi các nguồn từ trường bên ngoài. Điều này khiến quá trình nội suy lại cơ thể người trở nên không chính xác trong điều kiện môi trường sử dụng có nhiều nguồn từ trường mạnh như: khung kim loại, máy tính, ... Trong bài báo này, chúng tôi đề xuất một mô hình học máy cho phép dự đoán các tư thế thân trên của cơ thể người, từ 03 đầu vào ổn định (đầu, bàn tay phải, bàn tay trái), qua đó, giảm phụ thuộc vào các cảm biến IMU.

**Từ khóa:** Cảm biến góc quay kết hợp cảm biến gia tốc (Inertial Measurement Unit - IMU); Decision Tree Regression (DTR).

### 1. ĐẶT VẤN ĐỀ

Cùng với sự phát triển của Cách mạng Công nghiệp 4.0, công nghệ thực tại ảo đang ngày càng giữ một vị trí quan trọng. Công nghệ thực tại ảo, cho phép tạo ra những trải nghiệm mô phỏng hết sức chân thực, qua đó nâng cao đáng kể trải nghiệm của người dùng. Bộ áo cảm biến cơ thể (Body motion tracking system) là một trong số đó. Thiết bị này, khi mặc lên người sẽ cho phép cảm biến chính xác các chuyển động của cơ thể người, qua đó đồng bộ về tư thế, vị trí của cơ thể người trong không gian.



**Hình 1.** (a) 03 vị trí lắp cảm biến Vive Tracker;  
(b) Các vị trí lắp cảm biến IMU đối với một thiết bị theo dõi chuyển động cơ thể thông thường.

Cảm biến IMU không những cho phép tạo ra bộ áo với chi phí thấp, độ nhạy, tốc độ xử lý cao mà còn rất nhỏ gọn, tiêu thụ điện năng thấp. Tuy nhiên, một nhược điểm lớn của IMU là độ chính xác dễ bị ảnh hưởng bởi từ trường nhiễu loạn [1, 2]. Điều này khiến mô hình người thu về

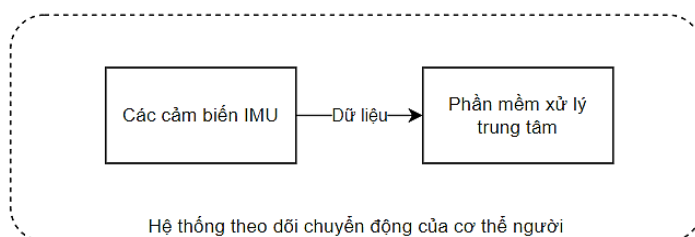
từ bộ áo cảm biến bị xoắn vặn, không chính xác. Đây là vấn đề nghiêm trọng, do nó ảnh hưởng trực tiếp tới tính khả dụng của cả bộ áo cảm biến cơ thể. Để giải quyết vấn đề này, người ta có thể tiếp cận theo hai phương pháp. Phương pháp vật lý, dựa trên việc sử dụng thêm các loại cảm biến khác, hoặc tìm cách hạn chế sự ảnh hưởng của từ trường nhiễu loạn thông qua bọc các nguồn từ trường bằng vật liệu cách từ. Phương pháp dựa trên giải thuật, là phương pháp tìm cách hiệu chỉnh kết quả đầu ra của cảm biến IMU thông qua mô hình, thuật toán. Trên thực tế, cả hai phương pháp trên đều đang được tiến hành nghiên cứu và tìm thấy tính ứng dụng trong những bài toán khác nhau.

Trong bài báo này, chúng tôi đề xuất sử dụng một phương pháp dựa trên học máy, cho phép hiệu chỉnh lại kết quả đầu ra của bộ áo cảm biến cơ thể dựa trên IMU. Mô hình học máy nhận 03 đầu vào ổn định (không bị nhiễu bởi từ trường) là đầu, bàn tay phải, bàn tay trái, từ đó dự đoán ra vị trí, góc nghiêng của 07 bộ phận thuộc thân trên của người (hai bên vai, cánh tay trên, cánh tay dưới và lưng). 03 đầu vào ổn định được thu từ thiết bị 03 cảm biến Vive Tracker (hình 1). Các cảm biến này cho đầu ra tương tự như IMU nhưng hoạt động dựa trên công nghệ hồng ngoại, do đó, không bị ảnh hưởng bởi từ trường.

## 2. CÁC NGHIÊN CỨU LIÊN QUAN

### 2.1. Tổng quan việc ứng dụng cảm biến IMU trong theo dõi chuyển động của cơ thể người

Các hệ thống theo dõi chuyển động cơ thể người dựa trên IMU thường được cấu tạo gồm 02 thành phần chính (hình 1): các cảm biến IMU và phần mềm xử lý trung tâm. Trong đó, các cảm biến có nhiệm vụ thu nhận chuyển động, còn phần mềm xử lý trung tâm tiến hành nhận dữ liệu và xử lý nội suy tư thế, chuyển động của cơ thể người.



**Hình 2.** Các thành phần của hệ thống theo dõi chuyển động của cơ thể người.

Ứng dụng khả năng xác định vị trí, hướng và tốc độ trong không gian 3 chiều của IMU. Các bộ cảm biến chuyển động được xây dựng thường bao gồm từ 8 đến 19 cảm biến IMU. Mỗi cảm biến được gắn tại các vị trí khác nhau trên cơ thể người. Khi cơ thể người chuyển động, các cảm biến IMU này sẽ thu nhận lại các chuyển động, từ đó, cho phép phần mềm đầu cuối nội suy ra tư thế, vị trí của cơ thể người trong không gian 3 chiều. Để việc nội suy được thuận tiện nhanh chóng, dữ liệu từ cảm biến IMU thường được tiền xử lý và trả về đầu ra ở dạng 6 bậc tự do gồm: vị trí và góc nghiêng theo 3 trục  $x, y, z$ .

### 2.2 Các nghiên cứu hiệu chỉnh nhiễu cho IMU

Vấn đề hiệu chỉnh, khử nhiễu cho cảm biến IMU thông qua giải thuật đã và đang được nhiều nhóm nghiên cứu trên thế giới quan tâm, thực hiện. Một số phương pháp điển hình có thể kể đến là xây dựng các giải thuật hiệu chỉnh trên cơ sở kết hợp với dữ liệu từ một số loại cảm biến hỗ trợ khác. Năm 2018, Solin và các cộng sự đã công bố một giải pháp lọc nhiễu [3,4], trong đó, nhóm tác giả sử dụng thêm dữ liệu video được thu thập từ camera, đồng thời với dữ liệu IMU. Các dữ liệu này sau đó được hiệu chỉnh bằng bộ lọc Kalman mở rộng (Extended Kalman Filter - EKF). Cũng trong những công bố này, nhóm tác giả cũng đề xuất thêm một số kỹ thuật cho phép hỗ trợ cho quá trình hiệu chỉnh như: vị trí cố định (là những vị trí mà tác giả cho rằng có dữ liệu ổn định theo thời gian), phương pháp cập nhật tốc độ của IMU. Cũng sử dụng kết hợp với dữ liệu camera, tuy nhiên, các công trình [5, 6] tiến hành ước lượng giá trị 6 bậc tự do của chuyển động,

thông qua các mạng nơ-ron học sâu. Các phương pháp này, tuy đạt hiệu quả nhưng đòi hỏi camera phải được cài đặt cố định và không bị che khuất trong cả quá trình vận hành. Bên cạnh đó, các giải pháp sử dụng mạng nơ-ron học sâu đối với dữ liệu video thời gian thực cũng đòi hỏi một năng lực tính toán cao.

Một số nghiên cứu khác dựa trên việc khai thác các giai đoạn ổn định của IMU. Bằng việc xác định các giai đoạn trong đó IMU là cố định, người ta có thể đặt các ngưỡng thay đổi, từ đó, hiệu chỉnh lại giá trị đầu ra của IMU. Phương pháp này thường được áp dụng trong bài toán theo dõi chuyển động của cơ thể người, do đặc trưng vận động của cơ thể người. Ý tưởng xây dựng các giải thuật cho phép xác định các giai đoạn IMU đạt tốc độ bằng 0 (zero-velocity), qua đó thiết lập ngưỡng đầu ra đã được nghiên cứu và công bố trong [7,8]. Tuy nhiên hiệu quả của phương pháp này phụ thuộc lớn vào ứng dụng cụ thể, cũng như chất lượng của từng loại IMU [9].

Cùng với đó, các phương pháp hiệu chỉnh dựa trên học máy cũng rất được quan tâm nghiên cứu. Trong năm 2019 [10], Lima và cộng sự đã đề xuất một kiến trúc mạng nơ-ron tích chập (Convolutional Neural Network – CNN) kết hợp với mạng bộ nhớ dài-ngắn hai chiều (Bidirectional Long Short-Term Memory - BLSTM) để dự đoán dữ liệu đầu ra của IMU.

Có thể thấy, bài toán hiệu chỉnh nhiều cảm biến IMU đã thu hút được rất nhiều sự quan tâm, nghiên cứu từ cộng đồng khoa học trên thế giới. Tuy nhiên, các giải thuật đề xuất hoặc chỉ quan tâm đến hiệu chỉnh, lọc nhiễu IMU mà chưa đặt vấn đề bày trong bối cảnh thu nhận chuyển động của cơ thể người; hoặc sử dụng các kiến trúc phức tạp, khiến mô hình có tính ứng dụng không cao.

### 3. PHƯƠNG PHÁP ĐỀ XUẤT

#### 3.1. Mô hình bài toán

Gọi đầu ra 6 bậc tự do của một cảm biến IMU  $x_i$ , cảm biến gắn trên vị trí  $i$ , tại thời điểm  $t$  là vector:

$$x_i^t = [px_i^t, py_i^t, pz_i^t, rx_i^t, ry_i^t, rz_i^t] \quad (1)$$

Trong đó:  $px, py, pz, rx, ry, rz$  lần lượt là vị trí và góc nghiêng theo 3 trục x, y, z của cảm biến IMU trong không gian 3 chiều. Như vậy, tư thế của cơ thể người trong không gian 3 chiều tại thời điểm  $t$  được biểu diễn bằng:

$$H^t = \{x_i^t \mid \forall i \in I\} \quad (2)$$

Trong đó,  $I$  là tập các cảm biến gắn trên cơ thể người. Do chuyển động của cơ thể người là có quy luật, do đó, các  $x_i^t$  không biến đổi ngẫu nhiên mà luôn tồn tại các ràng buộc lẫn nhau. Các ràng buộc này được biểu diễn bằng phân phối:

$$D_i = P_i(x_i^t \mid x_j^t \in X) \quad (3)$$

Do đặc trưng di chuyển của cơ thể người,  $x_i$  không phụ thuộc vào các cảm biến  $x_j$  như nhau, mà sẽ phụ thuộc chính vào một số cảm biến nhất định. Trong bài báo này, chúng tôi giả thuyết rằng, các cảm biến đặt ở thân trên sẽ phụ thuộc chính vào 03 cảm biến ở vị trí 2 bàn tay  $x_l$  (bàn tay trái),  $x_r$  (bàn tay phải) và  $x_h$  (đầu). Như vậy, công thức (3) đối với các cảm biến ở nửa thân trên có thể được biểu diễn như sau:

$$D_i \approx D_i^* = P_i(x_i^t \mid x_l^t, x_r^t, x_h^t) \quad (4)$$

$$D = \{D_i^* \mid \forall i \in \bar{I}\} \quad (5)$$

Trong đó,  $D_i^*$  là phân phối cho cảm biến  $i$  theo 03 cảm biến  $x_l, x_r$  và  $x_h$ .  $D$  là phân phối của của toàn bộ 7 cảm biến thuộc thân trên với điều kiện  $x_l, x_r$  và  $x_h$ . Để giải bài toán này, chúng ta cần tìm  $D$ , tuy nhiên việc tính toán  $D$  một cách trực tiếp là hết sức khó khăn, do đó, chúng tôi tiến hành ước lượng  $D$  thông qua các mô hình học máy  $F$ .

$$D \approx F(x_l, x_r, x_h) \quad (6)$$

Bằng cách mô hình hóa trên, có thể thấy, bài toán trở thành vấn đề hồi quy trong học máy. Trong đó người ta có thể áp dụng một lớp các mô hình học máy cho bài toán hồi quy để tìm  $F$ . Trong bài toán này, chúng tôi lựa chọn mô hình Hồi quy cây quyết định (Decision Tree Regression - DTR) để ước lượng  $F$ .

### 3.2. Xây dựng dữ liệu

Dữ liệu được tiến hành thu thập thông qua bộ áo cảm biến cơ thể Perception Neuron [11] và bộ 03 cảm biến Vive Tracker. Trong đó, các cảm biến Vive Tracker được đặt ở vị trí bàn tay phải, trái và đầu. Mỗi vị trí cảm biến trên cơ thể sẽ được thu thập 06 giá trị gồm: vị trí và hướng theo 3 trục  $x, y, z$ . Dữ liệu được thu thập với tần số lấy mẫu 10Hz. Trong thời gian thu dữ liệu, người mặc bộ áo cảm biến cơ thể tiến hành thao tác đi bộ, trong điều kiện từ trường ổn định, không tiếp xúc gần với các nguồn từ trường mạnh. Kết quả thu được 4.432 mẫu dữ liệu.

Trong đó, mỗi mẫu dữ liệu gồm 10 vector 6 chiều. Trong đó, 3 vector (thu từ 3 cảm biến ở vị trí tay phải, trái và đầu) được sử dụng làm đầu vào và 7 vector còn lại được sử dụng là nhãn (đầu ra) của mô hình. Dữ liệu được chia về 2 tập con: tập huấn luyện (chiếm 80% số mẫu, tương đương 3.545 mẫu), tập kiểm tra (887 mẫu).

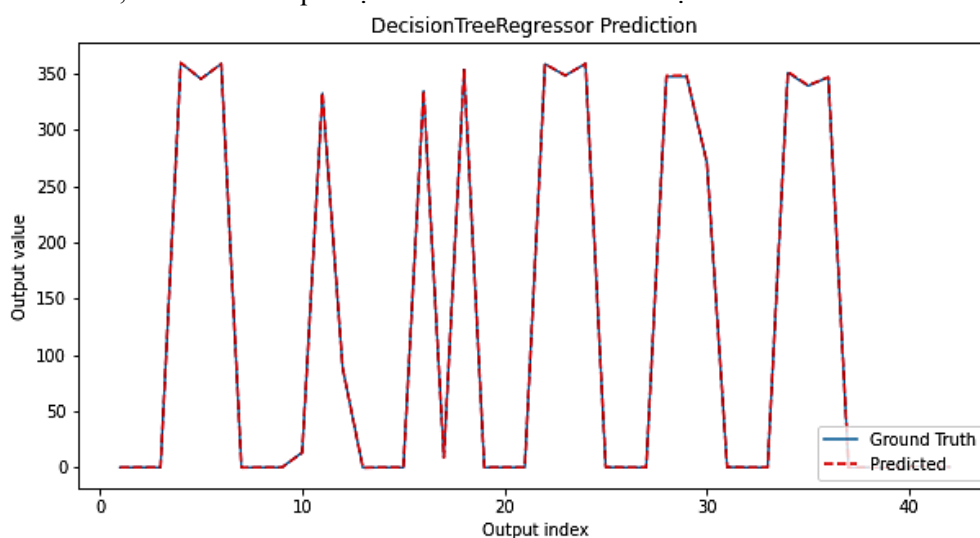
## 4. KẾT QUẢ VÀ THẢO LUẬN

Sau khi mô hình đề xuất được huấn luyện xong, chúng tôi tiến hành so sánh kết quả đạt được trên tập dữ liệu kiểm tra với 02 mô hình khác dựa trên kiến trúc MLP (Multi-layer Perceptron) và CNN (Convolutional Neural Network) theo bộ tham số: MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), RMSLE (Root Mean Squared Log Error). Kết quả cho thấy mô hình DTR có sai số nhỏ nhất.

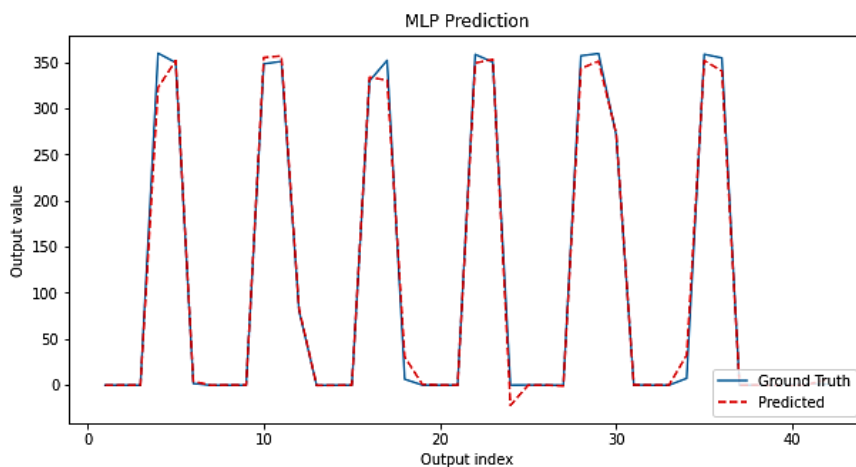
**Bảng 1.** Kết quả đánh giá các mô hình trên tập dữ liệu kiểm tra.

	MAE	RMSE	RMSLE
<b>DTR</b>	<b>2,3</b>	<b>27.96</b>	<b>3.33</b>
<b>MLP</b>	9.64	36.6	3.6
<b>CNN</b>	15.7	40.0	3.69

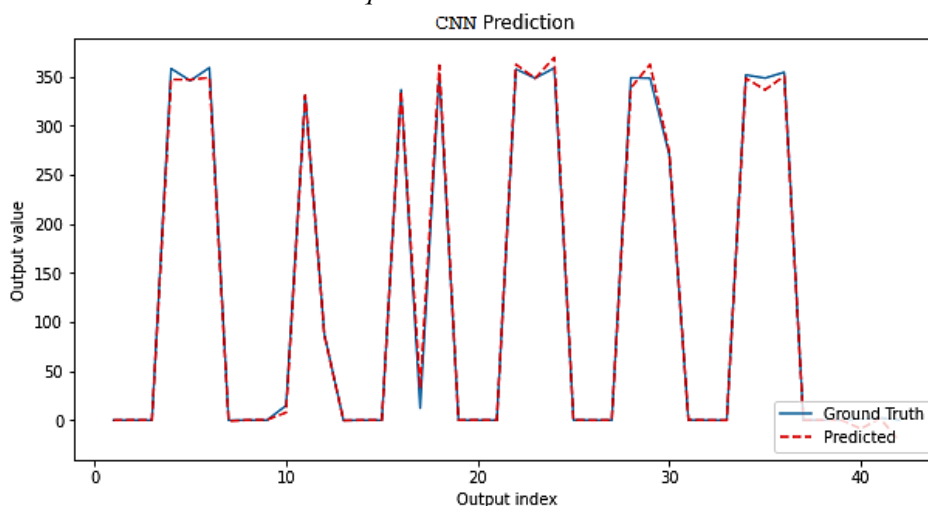
Mô hình hóa, biểu diễn kết quả dự đoán của các mô hình được cho như sau:



**Hình 3.** Biểu diễn kết quả đầu ra của mô hình DTR so với nhãn.



Hình 4. Biểu diễn kết quả đầu ra của mô hình MLP so với nhãn.



Hình 5. Biểu diễn kết quả đầu ra của mô hình CNN so với nhãn.

Kiến trúc MLP và CNN được lựa chọn để so sánh do đây là 02 kiến trúc được sử dụng rất phổ biến hiện nay. Trong đó, MLP là kiến trúc đặc biệt thông dụng đối với bài toán hồi quy với dữ liệu dạng bảng.

Dữ liệu IMU thu được có miền giá trị từ 0 – 359 (tương ứng với góc quay), do đó, dữ liệu thay đổi đột ngột khi biến thiên qua giá trị 359. DTR cho kết quả tốt nhất do giải thuật này không đòi hỏi dữ liệu phải chuẩn hóa trước khi huấn luyện. Chính vì vậy, DTR dễ dàng mô hình hóa được tính chất này. Bên cạnh việc có độ sai số thấp hơn, việc ứng dụng mô hình DTR vào bài toán hiệu chỉnh dữ liệu của cảm biến IMU có một số ưu điểm sau: thứ nhất, mô hình này có tốc độ tính toán nhanh, chiếm dụng tài nguyên thấp; thứ hai, việc cài đặt, áp dụng mô hình DTR là đơn giản và thuận tiện hơn so với các mô hình dựa trên mạng nơ-ron.

## 5. KẾT LUẬN

Trong bài báo, chúng tôi đã đề xuất phương pháp và mô hình dự đoán tư thế thân trên của cơ thể người chỉ với 03 cảm biến, so với bộ thiết bị gồm 10 cảm biến IMU. Kết quả so sánh các kết quả dự đoán so với nhãn cho sai số MAE xấp xỉ 2,3. Phương pháp này nếu được ứng dụng trong thực tiễn, không những có thể làm giảm số cảm biến trong bài toán theo dõi chuyển động của cơ thể người, đồng thời có thể loại bỏ các ảnh hưởng của nhiễu từ trường.

Tuy nhiên, tại thời điểm hiện tại, mô hình này mới chỉ phục vụ dự đoán cho các cảm biến IMU đặt ở thân trên của người. Trong tương lai, nhóm nghiên cứu sẽ tiếp tục phát triển mô hình dự đoán cho các cảm biến đặt ở những vị trí còn lại.

*Lời cảm ơn:* Nghiên cứu này được tài trợ bởi đề tài cấp Bộ Quốc Phòng, mã số: KC.BM.04.

### TÀI LIỆU THAM KHẢO

- [1]. Robert-Lachaine, Xavier & Mecheri, Hakim & Larue, Christian & Plamondon, Andre. “Effect of local magnetic field disturbances on inertial measurement units accuracy”. Applied Ergonomics. 63, (2017). 123-132. 10.1016/j.apergo.2017.04.011.
- [2]. Groves, P.D. “Navigation using inertial sensors [Tutorial]”. IEEE Aerosp. Electron. Syst. Mag., 30, 42–69, (2015).
- [3]. Solin, A.; Cortes, S.; Rahtu, E.; Kannala, J. “PIVO: Probabilistic inertial-visual odometry for occlusion-robust navigation”. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, pp. 616–625, (2018).
- [4]. Solin, A.; Cortes, S.; Rahtu, E.; Kannala, J. “Inertial odometry on handheld smartphones”. In Proceedings of the 2018 21st International Conference on Information Fusion (FUSION), Cambridge, UK, pp. 1–5, (2018).
- [5]. Clark, R.; Wang, S.; Wen, H.; Markham, A.; Trigoni, N. “Vinet: Visual-inertial odometry as a sequence-to-sequence learning problem”. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, (2017).
- [6]. Chen, C.; Rosa, S.; Miao, Y.; Lu, C.X.; Wu, W.; Markham, A.; Trigoni, N. “Selective sensor fusion for neural visual-inertial odometry”. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp. 10542–10551, (2019).
- [7]. Li, X.F.; Mao, Y.L.; Xie, L.; Chen, J.B.; Song, C.L. “Applications of zero-velocity detector and Kalman filter in zero velocity update for inertial navigation system”. In Proceedings of the 2014 IEEE Chinese Guidance, Navigation and Control Conference, Yantai, China, pp. 1760–1763, (2014).
- [8]. Li, L.; Pan, Y.; Lee, J.K.; Ren, C.; Liu, Y.; Grejner-Brzezinska, D.A.; Toth, C.K. “Cart-mounted geolocation system for unexploded ordnance with adaptive ZUPT assistance”. IEEE Trans. Instrum. Meas. 61, 974–979, (2012).
- [9]. Yao, Y.; Xu, X.; Xu, X. “An IMM-aided ZUPT methodology for an INS/DVL integrated navigation system”. Sensors, 17, 2030, (2017).
- [10]. Silva do Monte Lima, J.P.; Uchiyama, H.; Taniguchi, R.I. “End-to-End Learning Framework for IMU-Based 6-DOF Odometry”. Sensors, 19, 3777, (2019).
- [11]. <https://neuronmocap.com/pages/perception-neuron-3>.

### ABSTRACT

#### **A machine learning-based method in body movement tracking with a small number of sensors**

*Most of the current body sensing devices are composed of inertial measurement units (IMUs). The IMU sensors are placed at a different positions on the human body and sense their position, rotation, and tilt angle in space, thereby interpolating the movement of parts and the entire human body. Although IMU sensors have high accuracy and fast processing speed, they suffer from a major limitation of being susceptible to external magnetic field sources. This makes the process of re-interpolating the human body become inaccurate in an environment where many strong magnetic fields exist, such as metal frames and computers. In this paper, we propose a model to predict the postures of the upper human body from 03 stable inputs (head, right hand, left hand), thereby reducing the usage of IMU sensors.*

**Keywords:** Inertial Measurement Unit - IMU; Decision Tree Regression (DTR).