

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**



**BÁO CÁO HÀNG TUẦN**  
**HỌC PHẦN: THỰC TẬP CƠ SỞ**

**ĐỀ TÀI:**  
**NGHIÊN CỨU ƯỚC LƯỢNG KHOẢNG CÁCH**  
**BẰNG CAMERA 2D**

Giảng viên hướng dẫn: TS. Kim Ngọc Bách

Sinh viên thực hiện:

B22DCCN634

Trần Hữu Phúc

**8-15/03/2025**

## **A. MỞ ĐẦU**

### **1. Lý do chọn đề tài**

Trong kỷ nguyên số hóa hiện nay, sự phát triển mạnh mẽ của công nghệ thông tin đã tạo ra một nhu cầu cấp thiết về việc tối ưu hóa tương tác giữa con người và máy móc, đặc biệt trong các lĩnh vực cần xác định khoảng cách chính xác trong không gian 3D. Hiện tại, việc đo lường khoảng cách chủ yếu dựa vào các cảm biến chuyên dụng như cảm biến siêu âm, laser hay cảm biến độ sâu, khiến chi phí triển khai cao và hạn chế khả năng ứng dụng rộng rãi.

Qua quá trình tìm hiểu, em nhận thấy các công trình nghiên cứu đã công bố còn một số hạn chế đáng kể. Một số thách thức được đưa ra như: độ chính xác bị ảnh hưởng bởi điều kiện ánh sáng, góc nhìn camera, sự chuyển động của đối tượng gây mờ ảnh và đặc biệt là việc chuyển đổi từ thông tin 2D sang không gian 3D một cách chính xác. Từ những thách thức này, em quyết định chọn đề tài "Nghiên cứu ước lượng khoảng cách bằng camera 2D" để nghiên cứu sâu hơn, với mong muốn đóng góp vào việc phát triển một phương pháp ước lượng khoảng cách hiệu quả, chính xác, tiết kiệm chi phí và tài nguyên, góp phần thúc đẩy sự phát triển của các ứng dụng tương tác thông minh trong tương lai.

### **2. Mục tiêu nghiên cứu**

Nghiên cứu và ứng dụng hồi quy phi tuyến tính vào bài toán ước lượng khoảng cách từ bàn tay đến camera 2D.

Xây dựng được mô hình có khả năng ước lượng khoảng cách bàn tay với sai số thấp. Từ đó đánh giá và so sánh độ chính xác của các phương pháp ước lượng khoảng cách khác nhau trong các điều kiện môi trường, ánh sáng khác nhau.

### **3. Nội dung nghiên cứu**

Nội dung 1: Nghiên cứu các kỹ thuật phát hiện bàn tay và các kỹ thuật ước lượng khoảng cách từ ảnh RGB.

Nội dung 2: Nghiên cứu phát triển mô hình ước lượng khoảng cách bàn tay dựa trên khái niệm hồi quy phi tuyến tính.

### **4. Đối tượng, phạm vi, phương pháp nghiên cứu.**

- Đối tượng nghiên cứu:
  - + Bàn tay người.
  - + Khung xương 3D của bàn tay.
  - + Nonlinear regression (hồi quy phi tuyến tính).
- Phạm vi nghiên cứu:
  - + Phát hiện bàn tay trên ảnh.
  - + Ước lượng khoảng cách bàn tay tới camera 2D với thuật toán đề xuất.
- Những phương pháp nghiên cứu:
  - + Phương pháp nghiên cứu mô hình hóa: Nghiên cứu và xây dựng mô hình hồi quy phi tuyến tính cho bài toán ước lượng khoảng cách bàn tay từ ảnh 2D. Thiết kế thuật toán chuyển đổi từ tọa độ 2D sang ước lượng khoảng cách 3D.
  - + Phương pháp nghiên cứu thực nghiệm: Cài đặt thực tế mô hình, thuật toán đề xuất sử dụng cho bài toán ước lượng khoảng cách bàn tay từ video.
  - + Phương pháp nghiên cứu tham khảo ý kiến chuyên gia: Đánh giá tính khả thi của các phương pháp ước lượng khoảng cách tay bằng hồi quy phi tuyến tính. Triển khai ý tưởng, cài đặt thực nghiệm, phân tích, đánh giá kết quả.
  - + Phương pháp nghiên cứu điều tra, khảo sát: Tổng hợp các nghiên cứu liên quan đến ước lượng khoảng cách 3D từ ảnh 2D. Đánh giá ưu nhược điểm của các phương pháp hiện có.

## **B. BÁO CÁO TIẾN ĐỘ**

### **1.1. Nghiên cứu khái niệm cơ bản của học sâu**

Học sâu (Deep Learning – DL) là một trong những phương pháp học máy có sử dụng nhiều lớp biến đổi phi tuyến trên dữ liệu đầu vào từ đó trích xuất được các đặc trưng của dữ liệu. Trong khi học, dữ liệu được xử lý qua nhiều lớp với các mức độ khác nhau. Dữ liệu có gán nhãn và đủ lớn thường được sử dụng để huấn luyện trong DL.

Deep Learning là một tập con của Học máy (Machine Learning - ML), mặt khác là một chức năng của trí tuệ nhân tạo (Artificial Intelligence - AI). Thuật ngữ AI thường được dùng để đề cập đến các kỹ thuật sử dụng máy tính “học” các hành vi của con người trong đó ML lại là ứng cử viên sử dụng một tập các thuật toán dùng dữ liệu đầu vào để huấn luyện và làm cho những yêu cầu trở thành khả thi.

Deep Learning học cách sử dụng giống cấu trúc của bộ não con người. Tương tự như cách con người đưa ra quyết định, các thuật toán DL học cách đưa ra kết luận dựa trên việc phân tích dữ liệu với một cấu trúc logic nhất định. DL làm được điều này bằng cách ứng dụng mạng thần kinh, tức là sử dụng cấu trúc nhiều lớp của các thuật toán.

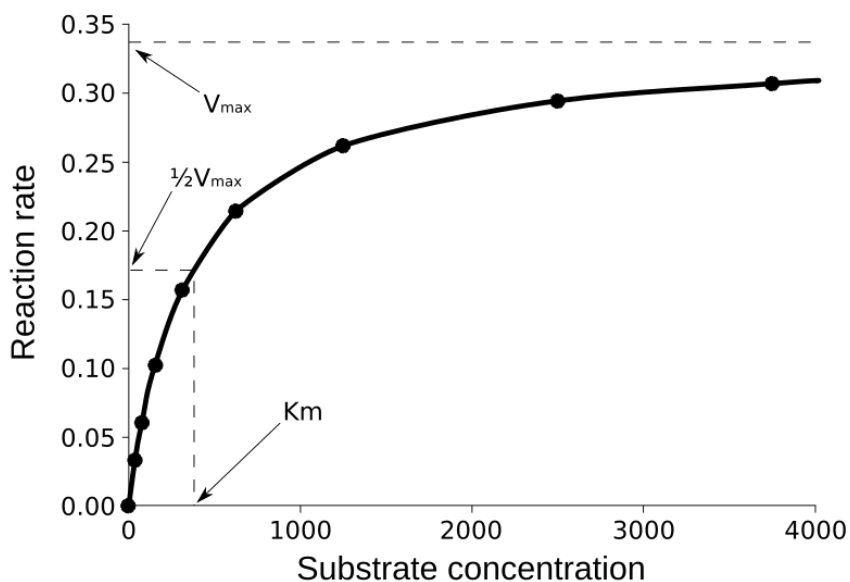
Mạng nơron nhân tạo mô phỏng cấu trúc và cách hoạt động của não người, sử dụng các lớp riêng lẻ như bộ lọc để phân tích thông tin từ tổng thể đến chi tiết, giúp phát hiện và đưa ra kết quả chính xác hơn. So với ML, DL không cần trích xuất tính năng, trong khi các thuật toán truyền thống như Cây quyết định, Support Vector Machine (SVM) và Naïve Bayes đòi hỏi bước này. DL cho phép các mạng nơron tự học cách biểu diễn dữ liệu thô, nén và trừu tượng hóa dữ liệu qua nhiều lớp, từ đó tạo ra kết quả như phân loại dữ liệu thành các lớp khác nhau.

Các mô hình DL tự động thực hiện quá trình trích xuất tính năng, giảm thiểu hoặc loại bỏ nhu cầu can thiệp thủ công. Trong khi ML yêu cầu con người xác định và cung cấp các đặc điểm cụ thể của đối tượng để mô hình phân loại, DL tự động nhận diện các đặc điểm này từ dữ liệu đầu vào. Một ưu điểm quan trọng khác của DL là khả năng xử lý lượng dữ liệu khổng lồ, với khoảng 2,5 tỷ tỷ byte dữ liệu được tạo ra mỗi ngày từ các nguồn như mạng xã hội, phương thức liên lạc và dịch vụ điện tử, cung cấp nguồn tài nguyên lớn cho sự phát triển của DL.

## 1.2. Nghiên cứu khái niệm cơ bản của hồi quy phi tuyến tính và mô hình bình phương của hồi quy phi tuyến tính

### 1.2.1. Hồi quy phi tuyến tính là gì?

Hồi quy phi tuyến tính (nonlinear regression) là một phương pháp phân tích thống kê dùng để mô hình hóa mối quan hệ giữa một biến phụ thuộc  $y$  và một hoặc nhiều biến độc lập  $x$ , trong đó quan hệ giữa chúng không phải là một hàm tuyến tính. Thay vào đó, mối quan hệ được thể hiện qua các hàm phi tuyến, ví dụ như lũy thừa, logarit, hàm mũ, hoặc các hàm khác.



Hồi quy phi tuyến tính thường được sử dụng khi dữ liệu có xu hướng không thể được biểu diễn chính xác bởi mô hình tuyến tính.

Dạng tổng quát của hồi quy phi tuyến tính có thể viết như sau:

$$y = f(x, \theta) + \epsilon$$

$f(x, \theta)$ : là một hàm phi tuyến tính của các biến độc lập  $x$ , tham số  $\theta$  là các hệ số cần ước lượng.

$\epsilon$ : là nhiễu hoặc sai số ngẫu nhiên.

Có nhiều phương pháp hồi quy khác nhau có thể được áp dụng để mô hình hóa tập dữ liệu, chẳng hạn như hồi quy bậc hai, hồi quy bậc ba và các dạng hồi quy bậc cao hơn, tùy thuộc vào yêu cầu cụ thể của bài toán.

Những giả định trong hồi quy phi tuyến tính tương tự như trong hồi quy tuyến tính nhưng có thể mang những ý nghĩa tinh chỉnh hơn do tính phi tuyến của mô hình. Dưới đây là các giả định chính trong hồi quy phi tuyến tính:

- Hình thức hàm số (Functional Form): Mô hình phi tuyến tính được chọn phải phản ánh chính xác mối quan hệ thực tế giữa biến phụ thuộc và biến độc lập.
- Tính độc lập (Independence): Các quan sát được giả định là độc lập với nhau.
- Tính đồng nhất phương sai (Homoscedasticity): Phương sai của phần dư (chênh lệch giữa giá trị quan sát và giá trị dự đoán) phải không đổi ở mọi mức của biến độc lập.
- Tính phân phối chuẩn (Normality): Phần dư được giả định tuân theo phân phối chuẩn.
- Đa cộng tuyến (Multicollinearity): Các biến độc lập không được hoàn toàn tương quan với nhau.

Trong học máy, hồi quy phi tuyến tính được phân thành hai loại chính:

- Hồi quy phi tuyến tính tham số (Parametric Nonlinear Regression): Loại này giả định rằng mối quan hệ giữa biến phụ thuộc và biến độc lập có thể được mô hình hóa bằng một hàm toán học cụ thể. Ví dụ, mối quan hệ giữa dân số và thời gian có thể được mô hình hóa bằng hàm số mũ.
- Hồi quy phi tuyến tính phi tham số (Non - parametric Nonlinear Regression): Loại này không giả định rằng mối quan hệ giữa biến phụ thuộc và biến độc lập có thể được mô hình hóa bằng một hàm toán học cụ thể. Thay vào đó, các thuật toán học máy được sử dụng để tìm hiểu mối quan hệ này từ dữ liệu.

### 1.2.2. Mô hình bình phương của hồi quy phi tuyến tính

Trong hồi quy phi tuyến tính, mô hình bình phương (quadratic model) là một loại mô hình đặc biệt trong đó mối quan hệ giữa biến phụ thuộc và biến độc lập là một đa thức bậc hai.

Dạng tổng quát của mô hình bình phương:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$$

Trong đó:

$y$ : biến phụ thuộc

$x$ : biến độc lập

$\beta_0, \beta_1, \beta_2$ : các tham số hồi quy

$\epsilon$ : nhiễu hoặc sai số ngẫu nhiên

Mô hình này phù hợp khi mối quan hệ giữa  $y$  và  $x$  không phải là đường thẳng mà có dạng parabol. Tùy vào giá trị của  $\beta_2$ , đồ thị của  $y$  theo  $x$  có thể lõm xuống ( $\beta_2 > 0$ ) hoặc lõm lên ( $\beta_2 < 0$ ).