Tập dữ liệu có tổng độ lớn 6h53m, được chia thành 3 phần: huấn luyện, kiểm thử và xác thực. Trong đó 80% được sử dụng cho huấn luyện và xác thực, 20% cho kiểm thử được chia ngẫu nhiên. Độ chính xác và đồ thị hàm mất mát. Độ chính xác cuối cùng của mô hình 98.20%.

Sau quá trình huấn luyện, sử dụng bộ chuyển đổi TensorFlow Lite để chuyển mô hình thành một mảng byte chứa thông tin của mô hình được huấn luyện. Kết quả mô hình chuyển đổi với dung lượng 185,624 bytes, xấp xỉ 186 KB, phù hợp để triển khai trên phần cứng ESP32-S3. ESP32-S3 sở hữu 384 KB ROM, 320 KB SRAM, và 4 MB bộ nhớ Flash. Tổng tài nguyên sử dụng của chương trình triển khai cuối cùng là 23 KB RAM (7 %) và 590 KB Flash (9 %). Thời gian thực thi của từng bước để thể hiện trong bảng sau:

|  |  |  |
| --- | --- | --- |
|  | Quá trình | Thời gian (ms) |
| 1 | Tiền xử lý dữ liệu (FFT) | 38 |
| 2 | Mô hình phân loại dữ liệu | 27 |
| 3 | Hiển thị kết quả | 5 |
|  | **Tổng** | 70 |

Phân tích mô hình học máy chạy trên phần cứng ESP32-S3 dựa trên thông số sau:

"PLATFORM: Espressif 32 (6.7.0) > Espressif ESP32-S3-DevKitC-1-N16R8V (16 MB QD, 8MB PSRAM)

HARDWARE: ESP32S3 240MHz, 320KB RAM, 16MB Flash

Current: 47 mA

RAM: [= ] 7.0% (used 22936 bytes from 327680 bytes)

Flash: [= ] 9.0% (used 589061 bytes from 6553600 bytes)

Model size = 185624 bytes

Predictions (DSP: 38 ms., Classification: 27 ms., Anomaly: 0 ms.)"

Dựa vào các thông số trên, đưa ra các đánh giá về mô hình phân loại cũng như thời gian xử lý có đáp ứng được trên các thiết nhúng có tài nguyên hạn chế; và xử lý dữ liệu thời gian thực

Với các thông số được cung cấp, có thể thấy rằng việc triển khai mô hình học máy trên phần cứng ESP32-S3 là hoàn toàn khả thi và hiệu quả. ESP32-S3 có bộ vi xử lý 240MHz, đủ mạnh để xử lý các tác vụ phức tạp. Với 320KB RAM, ứng dụng chỉ sử dụng 22936 bytes, tức là khoảng 7.0% tổng RAM, cho thấy sự tiết kiệm bộ nhớ, đảm bảo còn đủ không gian cho các tác vụ khác hoặc dữ liệu bổ sung. Dung lượng Flash là 16MB, với ứng dụng hiện tại chỉ sử dụng 589061 bytes, tương đương 9.0% tổng dung lượng, giúp còn lại rất nhiều dung lượng để lưu trữ thêm dữ liệu, mã nguồn hoặc các phiên bản mô hình khác.

Mô hình học máy có kích thước 185624 bytes, khá nhỏ gọn so với dung lượng bộ nhớ có sẵn, cho thấy tính tối ưu của mô hình này. Về thời gian dự đoán, việc xử lý tín hiệu số (DSP) mất 38 ms và quá trình phân loại mất 27 ms, trong khi xử lý anomaly detection mất 0 ms, cho thấy thời gian này không đáng kể hoặc không có xử lý anomaly detection trong mô hình hiện tại. Tổng thời gian xử lý cho một dự đoán là khoảng 65 ms, đủ nhanh để thực hiện khoảng 15 dự đoán mỗi giây, phù hợp cho các ứng dụng yêu cầu xử lý gần thời gian thực.

ESP32-S3 cho thấy khả năng quản lý và sử dụng tài nguyên hiệu quả, đặc biệt quan trọng đối với các thiết bị nhúng có tài nguyên hạn chế. Với lượng RAM và Flash còn dư, có thể dễ dàng mở rộng hoặc phức tạp hóa mô hình mà không gặp vấn đề về tài nguyên. Hệ thống này không chỉ đáp ứng tốt các yêu cầu xử lý dữ liệu thời gian thực mà còn có khả năng xử lý dữ liệu liên tục mà không bị gián đoạn. Điều này làm cho ESP32-S3 trở thành một lựa chọn lý tưởng cho các ứng dụng như nhận dạng giọng nói, phát hiện chuyển động, và các ứng dụng IoT khác. Với hiệu suất xử lý nhanh chóng và khả năng quản lý tài nguyên hiệu quả, ESP32-S3 là một giải pháp mạnh mẽ cho việc triển khai mô hình học máy trên các thiết bị nhúng.