

Sistema de reconhecimento de atividades humanas utilizando sensores de movimento vestíveis

Jesimon Barreto, Jonas Reis Jr, Julinei S Carvalho, Lucas Bispo, Lucas Chaves

Abstract—Reconhecer atividades humanas utilizando sensores vestíveis é uma tarefa aplicada em áreas de diferentes aplicações de sistemas como casas inteligentes, internet das coisas, interação humano computador. Dessa maneira, esse trabalho apresenta um projeto que utiliza dados de sensores de movimento vestíveis para o reconhecimento de atividades humanas com abordagem baseada em aprendizado profundo, além disso, esse modelo compara modelos da literatura em hardwares de baixo custo.

Reconhecimento de atividade humana, Sensores vestíveis, Aprendizado de máquina, Dispositivos móveis

I. INTRODUÇÃO

ALGORITMOS que identificam atividades humanas automaticamente têm sido um problema amplamente pesquisado há muito tempo. Consiste em, dado uma entrada, sendo dados capturados por sensor (e.g. câmeras [15], sensores infravermelho [22], sensores vestíveis (acelerômetros, giroscópio, magnetômetro) [8], entre outros sensores que capture movimentos de uma ou mais pessoas), a saída deve ser a ação (e.g. correr, caminhar, cozinhar, subir escada, atirar flechas, dirigir, entre outros) que a pessoa está executando. A importância desse tipo de solução nos últimos tempos está se tornando cada vez mais evidente devido a aplicações em casas inteligentes, internet das coisas, vigilância, interação humano computador, detecção de anomalias, monitoramento de saúde, entre outros [8].

Existem soluções que fazem uso de diferentes tipos de sensores para utilizar seus dados como entrada. Dessa forma, escolher qual sensor utilizar como entrada é fundamental, pois esse processo traz limitações relacionadas com ambiente, atividades a serem detectadas, tamanho do sistema e relativas à quantidade de pessoas que o sistema deve prever as atividades [4]. Por exemplo, para o caso onde se utiliza câmeras em ambientes abertos ou fechados e se tem acesso a dados de várias pessoas em uma só captura, uma limitação que deve ser considerada é a distância da câmera em relação as pessoas, que pode causar problemas como oclusão, dificuldade para captura de movimentos mais sutis[15], além de que, câmeras ficam fixas e por isso não conseguem acompanhar as pessoas. Por outro lado, utilizando-se de sensores com câmeras e sensores de captura de movimento, como kinect [5], também há limitações, relacionado ao sensor, que funciona somente em ambientes internos, e a limitação da distância [22]. Além desses sensores citados, existe os sensores de movimento vestíveis (e.g. acelerômetro, giroscópio, magnetômetro) presentes em smartphones e smartwatches. Esses sensores se destacam nas situações de ambientes pois como as pessoas os levam com

elas, é possível detectar atividades naturalmente em qualquer ambiente, mas ainda sim apresentam limitações referentes a atividades que são reconhecidas somente com imagens (e.g. assistir TV, diferenciar entre comer com garfo e comer com colher, entre outros)[4].

Esse trabalho foca em reconhecimento de atividades utilizando sensores de movimentos vestíveis como entrada. Cada um desses sensores disponibiliza três séries temporais (e.g. X,Y,Z) que representam a variação de cada eixo, para cada sensor equivalente. No caso do acelerômetro, representa a variação de aceleração, no caso de giroscópio indica variação de ângulo de rotação do dispositivo e o magnetômetro indica a posição do celular no espaço físico, dado medições de campo magnéticos presentes [18].

A série temporal criada por cada dimensão do sensor, é assim chamada pois é construída por valores medidos em um determinado intervalo de tempo, gerando, durante um determinado período, um comportamento de linha no tempo [23] como apresentado na figura 1. O modelo deve ter como entrada esses sinais e a saída deve ser a classificação da atividade ao qual o sinal se refere.

Os modelos da literatura mostram que atualmente o estado da arte é baseado em modelos de aprendizado profundo[4]. Nessa perspectiva, alguns problemas enfrentados para reconhecimento de atividade humana utilizando sensores de movimento são: a heterogeneidade, o desbalanceamento de classes e a extração de características [4]. A heterogeneidade refere-se às diferentes maneiras de se executar uma atividade, como a diferença de velocidade, intensidade ou quantidade de repetições. Por exemplo, para a atividade de pentear o cabelo, existem diferentes tipos e tamanhos de cabelo que são relativos a cada pessoa e que também têm diferenças na velocidade, no tempo e na forma de se executar a ação. A representação é muito pequena em relação a quantidade de maneiras que as ações possam ser executadas. Outro problema é o desbalanceamento de amostras, entre vários fatores para esse problema, o maior está relacionado com a dificuldade de capturar amostras de determinada atividade. Exemplo disso é que, caminhar é uma atividade fácil de ser capturar, enquanto que escovar os dentes é mais difícil. Por fim, extração de características é uma dificuldade atrelada a natureza do problema, pelo fato de ser uma série temporal, existem atividades que são bastante parecidas e isso faz com que seja difícil distinguir, são exemplos correr e caminhar, sentar e levantar, deitar e levantar, entre outras[4].

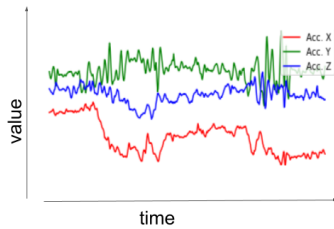


Fig. 1. Apresenta um exemplo de sinal temporal de 3 dimensões (X,Y,Z) de um acelerômetro de um dispositivo móvel que estava no pulso de uma pessoa.

A. Motivação

Vive-se, nos dias atuais, um avanço imenso do campo de inteligência artificial, aprendizado de máquinas e IoT, combinado a isso a utilização cada vez maior de sensores vestíveis faz com que o Reconhecimento de Atividades Humanas (RAH) seja um dos principais focos de pesquisa no desenvolvimento de tecnologias para auxiliar a sociedade. [8] Assim, a influência de dispositivos inteligentes e aplicativos de comunicação no dia a dia da população é cada vez mais presente. Na área da saúde, o RAH está sendo usado para a detecção prévia de doenças, no cuidado de idosos e no monitoramento de atividades físicas [6]. Na área de segurança, o RAH pode ter papel fundamental em garantir uma direção mais segura, como por exemplo alertando sobre excesso de sono ou consumo de álcool do motorista [8]. Além disso, a Indústria 4.0 vem crescendo cada ano mais e o uso de sensores e de RAH pode ser de muito auxílio em equipamentos IoT para automatização de casas inteligentes. Dessa forma, conseguir desenvolver um algoritmo que tenha uma alta precisão e que consiga ser usado em tempo real em dispositivos de baixo custo é uma tarefa bastante importante e vital na popularização dessa tecnologia que abre portas a diversas novas aplicações.

Os escopos que permeiam este projeto permitirão entender não apenas a correlação de RAH e IoT, mas também como esse entrelaçamento engendra impactos significativos em diversos setores da sociedade (e.g. economia, saúde, aspectos políticos e sociais). No que concerne aos aspectos econômicos, a empresa de consultoria McKinsey [12] prevê em 2025, a movimentação de até 11,1 trilhões de dólares de tecnologias decorrentes de nove pilares de IOT, sendo que o setor de Healthcare é um deles, conforme ilustra a figura abaixo.

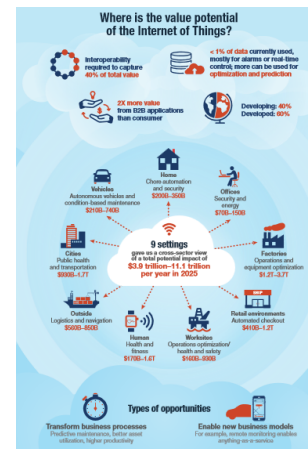


Fig. 2. Potencial econômico estimado dos impactos de IoT (e suas sub-áreas) em 2025 [12]

B. Originalidade da proposta

A originalidade da proposta está na análise de aplicabilidade na prática, que seria implementar o nosso modelo e os modelos de referência em um hardware de baixo custo e avaliá-los em termos de comportamento em nível de processamento, ocupação de memória e tempo de processamento do modelo. O que foi percebido na literatura, até então, é que os trabalhos encontrados se preocupam somente com modelo e taxa de acerto, porém esquecem do compromisso de que, na prática, os modelos devem ser utilizados em dispositivos móveis que apresentam limitação de processamento e memória.

C. Desafios

Como levantado na literatura, os desafios estão relacionados com o fato de arquiteturas de aprendizado profundo exigirem poder computacional, porém a aplicação desses métodos são para dispositivos de baixo custo[21]. Dessa maneira, desafios giram em torno de trabalhar com uma alta taxa de acertos a medida que é ponderado os gastos com processamento, memória e tempo de resposta.

D. Abrangência

O projeto é focado em desenvolver um sistema de reconhecimento de atividades humanas que seja robusto mas que seja aplicável em dispositivos vestíveis com pouca capacidade de processamento e sem dependência de internet.

E. Produção

É esperado que seja desenvolvido um modelo de classificação baseado em aprendizado profundo para reconhecer atividades humanas com foco em dispositivos low-power. Este modelo será desenvolvido com a aplicação dos seguintes conceitos: Internet das Coisas e *Edge Machine Learning*.

1) *Internet das Coisas*: A *Internet das Coisas* é o modo que os objetos físicos estão conectados e comunicando entre si e com o usuário final, por meio de um conjunto de sensores e utilização da internet como meio de comunicação. Tais objetos físicos podem ser qualquer coisa, seja um relógio, uma geladeira, televisão, cafeteira, balança, entre outros.

2) *Edge Machine Learning*: O conceito de *Edge Machine Learning* consiste em aplicar técnicas de machine learning só que com o processamento dos dados feito localmente, em servidores na rede local ou no próprio dispositivo que coletou os dados.

II. OBJETIVO GERAL E OBJETIVOS ESPECÍFICOS

Este projeto tem o objetivo principal de criar um método de reconhecimento de atividades humanas utilizando sensores de movimento vestíveis, que atuem em tempo real em dispositivos móveis e avaliar seus comportamento de tempo real em um dispositivo na prática.

Dado esse objetivo, foi avaliado e dividido entre os seguintes objetivos específicos:

- 1) Revisão de Literatura
Será realizado um trabalho para revisar o que há na literatura e modelos para comparação.
- 2) Pré-processamento dos dados
Os dados de treinamento serão explorados e preparados, de modo a facilitar a construção do modelo de reconhecimento de atividades.
- 3) Construção do Modelo
Desenvolvimento do modelo, esse processo pode levar muito tempo do projeto, está relacionado com implementações de modelos com dados disponíveis
- 4) Avaliação do Modelo
O modelo será avaliado com base nos sistemas de referência tirados da literatura. O passo anterior e esse passo poderão ser executados diversas vezes durante o projeto.
- 5) Documentação
A documentação do sistema será desenvolvida durante todo o tempo disponível para projeto.
- 6) Apresentação do sistema
É esperado que no último mês seja realizado a apresentação do sistema para os stakeholders.

III. RELEVÂNCIA DO PROJETO

O cenário tecnológico atual traz conceitos inovadores como ambientes inteligentes e internet das coisas que desenvolvem cada vez mais a interação humano computador.

Nesse sentido, o Reconhecimento de Atividades Humanas contribui para o desenvolvimento desse cenário classificando as atividades das pessoas e disponibilizando esses dados para diferentes aplicações que visam melhorar a qualidade de vida, garantir segurança, analisar performances profissionais, monitorar atividades do dia-a-dia e muitas outras. Porém, existem desafios na entrega de um modelo que atende os critérios do mercado em geral, podendo-se mencionar a necessidade de sistemas leves, ponderando os gastos com processamento

de grandes quantidades de dados, consumo de memória e consumo de energia, isso tudo com bom tempo de resposta e alta taxa de acerto.

A aplicabilidade do modelo proposto nesse projeto visa a implementação de RAH em um hardware de baixo custo, com otimização de processamento, uso de memória e energia, além de eficiência nos resultados. O modelo deve ser utilizado em dispositivos móveis, desde os mais populares que apresentam limitação de processamento e memória até os mais robustos com alto desempenho. Dessa forma, o modelo é aplicável nos sistemas das mais diferentes aplicações com o intuito de produzir um produto leve na operação e de fácil uso e aceitação por parte dos usuários, visando uma produção massiva no mercado.

IV. TRABALHOS RELACIONADOS

O reconhecimento de atividade humana utilizando dados de sensores (RAH é tratado na bibliografia como classificação de sinais temporais[11][4]. Porém, algumas diferenças na validação e na estruturação dos dados são essenciais. Jordão et al. [9] Sena et al.[19] e Sena et al. [20] propuseram entender o real significado dos dados. Além disso, Jordão et. al. [9] propôs uma forma de validação justa para comparação de modelos voltados para RAH.

Sena et. al. [20] propõe um método multimodal, nesse caso, utiliza os sensores presentes em um dispositivo móvel. Na arquitetura, é proposto um ensemble de redes convolucionais de kernels de tamanhos diferentes, caracterizado por analisar o sinal em várias escalas, capturando características suficientes para diferenciar atividades que são semelhantes. Rueda et. al. [17] segue essa mesma ideia. Porém, o custo computacional consumido por esse tipo de método cresce de acordo com a quantidade de sensores disponíveis.

Jiang and Yin [7] usou para pré-processamento short-time Discrete Fourier Transform [13], que gera uma representação baseada na frequência do sinal em uma seção local. Essa representação foi utilizada como uma imagem time-frequency-spectral usada para diferenciar atividades que tem comportamentos parecidos de movimento, porém acontecem em velocidades diferentes, por exemplo, caminhar e correr. Apresenta essa imagem para uma rede CNN 2D. A time-frequency-spectral não é suficiente para diferenciar atividades mais complexas.

Lu and Tong [11] propuseram como pré-processamento o modified recurrence plot, capaz de codificar o sinal de série temporal em uma imagem RGB. Com base nisso, utiliza essa imagem como entrada para uma arquitetura baseada em TinyResnet. O pré-processamento utiliza muito recurso computacional, a transformação.

Nosso trabalho apresenta um diferencial no processo de pré-processamento, pois utilizamos o máximo de representações do sinal possíveis. Dessa maneira, o modelo vê o sinal de diferentes modos, melhorando a capacidade de predição da rede. Outro fator é a avaliação do método em um hardware de baixo custo. Dos modelos apresentados nos trabalhos, a maior parte não apresenta a validação do modelo em um hardware de baixo custo. Considerando isso, uma contribuição será uma

metodologia para validação em hardware que considera o custo computacional consumido por cada trabalho.

V. REQUISITOS

Nesta seção serão apresentados os requisitos do projeto. Os requisitos foram divididos entre funcionais e não funcionais e são apresentados nas tabelas I e II.

TABLE I
REQUISITOS FUNCIONAIS

ID	Descrição
RF001	O modelo deverá funcionar offline
RF002	O modelo deverá classificar atividades em tempo real.
RF003	O modelo deverá conseguir classificar as mais diversas atividades do dia a dia, tais como caminhar, correr, dormir, cozinhar e etc.
RF004	As entradas do modelo devem ser obtidas através dos dados dos sensores vestíveis: Acelerômetro, Giroscópio e Magnetômetro
RF005	O modelo deve ter como saída a classificação da atividade realizada
RF006	O modelo deve permitir a adição de novas atividades
RF007	O modelo deve pre-processar os dados de entrada de modo a diminuir a complexidade computacional

Os requisitos funcionais que são necessários para o funcionamento básico do sistema estão citados e descritos na tabela I. Do mesmo modo que os requisitos não funcionais que são requisitos que estamos interessados em atingir mas que não são necessariamente básicos para o funcionamento do sistema modelado, os requisitos estão apresentados na tabela II.

TABLE II
REQUISITOS NÃO FUNCIONAIS

ID	Descrição	Sub-Categoria
RNF001	O modelo deverá obter uma boa acurácia comparável ao estado da arte.	Desempenho
RNF002	O modelo deverá ter um baixo consumo de energia.	Eficiência
RNF003	O modelo deverá ter no máximo 256Mb de tamanho.	Espaço
RNF004	O modelo deverá rodar em processadores com no mínimo 1GHz de Clock.	Portabilidade
RNF005	O modelo deverá rodar em computadores com no mínimo 512Mb de RAM	Portabilidade
RNF006	O modelo deverá rodar em dispositivos Linux	Produto
RNF007	O modelo deverá ser compatível com o Tensorflow	Produto
RNF008	O modelo deverá ser desenvolvido em Python	Produto
RNF009	O modelo deverá ter um recall comparável ao estado da arte	Robustez

VI. RESTRIÇÕES

Nesta seção é apresentado algumas restrições direcionadas a aplicação do sistema. As restrições do sistema proposto são:

- O sistema não deverá utilizar processamento em nuvem para realizar o reconhecimento das atividades.
- O sistema deverá ser capaz de rodar em dispositivos de baixo custo.

VII. METODOLOGIA

A figura 3 apresenta a metodologia proposta e indica que o dado foi capturado e, em seguida, separado nos 3 eixos. Esse dado é transformado em 3 ou 2 dessas representações e apresentado em uma arquitetura de *ensemble* de redes neurais profundas. A saída é a resposta esperada.

A seleção de quais sensores serão utilizados depende da disponibilidade do dispositivo móvel. Em sua maioria, os dispositivos no mercado apresentam acelerômetro, e alguns giroscópios, magnetômetro entre outros. Nosso método propõe utilizar quais estiverem disponíveis.

As representações que serão utilizadas dependem do tempo de transformação para cada sinal. Nosso método apresenta 3 transformações principais: a temporal, a plote de recorrência [11] e espectrograma [3]. Porém, pode haver casos que a transformação do plote de recorrência pode se demonstrar lento para usar, nesse caso prevemos não usá-lo.

Ademais, neste projeto, há a possibilidade de se usar certos modelos (e.g., modelo em Espiral e modelo em V) típicos em Engenharia de Sistemas baseada em modelos [10]. O uso desses modelos será relacionado a etapas específicas em fases de análises que serão desenvolvidas ao longo da etapa final deste projeto (etapa 4). Assim, o uso desses modelos será adequadamente explorado ao longo do desenvolvimento dessas etapas finais.

A. Representação Temporal

Essa representação é simples. Os valor de cada tempo do sinal 1D é tido como um pixel de uma imagem de 1 canal. Essa representação tem como dimensões $S \times T$, onde $S = 3 * \text{sensores}$ e T é o número capturas temporais por amostra. [17][19][20].

B. Plote de Recorrência Modificado

Uma matriz de recorrência é dado por $R(m, n)$, (m, n) é preenchido usando a equação 2 que por sua vez utiliza a equação 1. s^i é o vetor formado por i -ésimo e o $(i + 1)$ -ésimo valor temporal onde $i = 1 \dots n$, n é a janela temporal. A norma L2 é utilizada na equação 2. Por último é feita uma normalização na imagem utilizando a equação 3.

$$\text{sign}(m, n) = \begin{cases} -1, & \text{if } \frac{(s^m - s^n) \cdot \mathbf{v}}{\|s^m - s^n\| \cdot \|\mathbf{v}\|} < \cos(\frac{3\pi}{4}) \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

$$R_{m,n} = \text{sign}(m, n) \cdot \|s^i - s^j\| \quad (2)$$

$$I = \frac{M - \min(M)}{\max(M) - \min(M)} \quad (3)$$

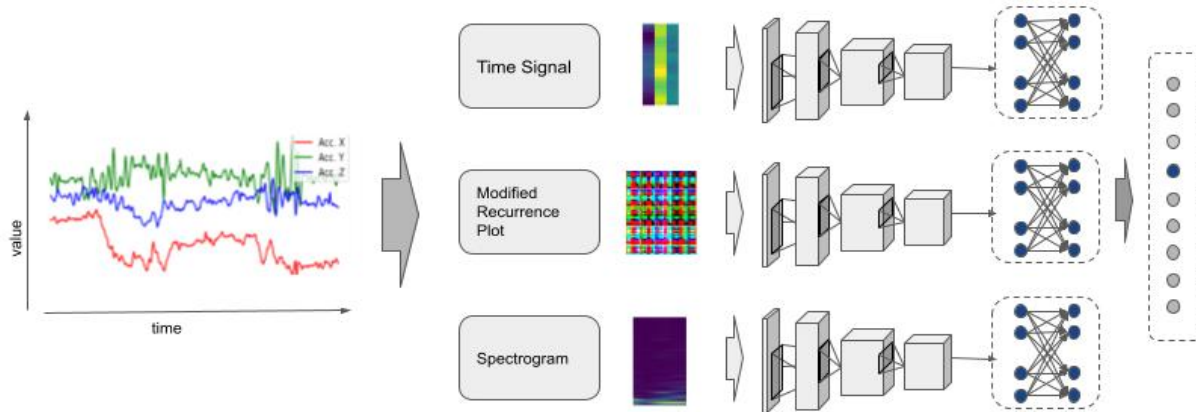


Fig. 3. Essa figura apresenta a metodologia deste trabalho. Entradas com visualizações diferentes dos dados são dados como entrada em um ensemble de redes neurais profundas. A saída é a predição do modelo.

C. Espectrograma

Espectrograma é a representação de séries temporais no domínio da frequência. Esse domínio é tido quando aplicada a transformada de Fourier. Essa representação é bastante usada para diferenciar a visualização de sinais [14]. Essas informações são utilizadas para diferenciar características pertencentes ao comportamento da frequência dos sinais.

D. Vantagens e Desvantagens do Modelo

Ao construir um modelo que considera dispositivos *low-power* e aplica os conceitos de *Edge Machine Learning* e Internet das Coisas, temos as seguintes vantagens:

- Menor consumo de processamento.
- Menor consumo de memória.
- Não necessita de acesso à internet, dado que os dados são processados localmente.
- Baixo consumo de energia, uma vez que não se tem um processador tão potente.
- Classificação em Tempo Real, dado que os dados são processados localmente.
- Maior privacidade e segurança de dados, dado que os dados são processados localmente.

Por outro lado, por se tratar de um dispositivo *low-power* e sem necessidade de acesso à internet, temos as seguintes desvantagens:

- Acurácia inferior à modelos mais robustos, uma vez que temos menor capacidade de processamento e precisamos de fazer alguns relaxamentos no modelo para garantir uma resposta em tempo real.
- Gerenciamento da Infraestrutura, dependendo do contexto, é necessário gerenciar a infraestrutura de suporte dos dispositivos que farão o uso desse modelo.
- Custo, apesar de ser um modelo focado em dispositivos *low-power*, ainda é necessário que esse dispositivo tenha uma capacidade razoável de processamento, o que encarece o dispositivo.

E. Avaliação de Desempenho em Hardware

A avaliação de desempenho em hardware segue um roteiro estabelecido visando a um andamento correto e padronizado. A premissa que se segue é o uso de um hardware de baixo custo comparável com o sistema ao qual os métodos terão sua aplicabilidade. Sena et. al. [16] propõe diferentes métricas para avaliação das quais definimos em nossa metodologia as seguintes:

- Porcentagem do processamento utilizado
- Porcentagem de memória RAM utilizada
- Tempo de processamento (conjunto de dados)
- Tamanho do arquivo em memória secundária
- Tempo de subir modelo (secundária para RAM)

Essas métricas definem o melhor modelo considerando um menor tempo de processamento dos dados utilizando-se da menor porcentagem do processamento possível. Assim como o processamento, a menor porcentagem de memória RAM utilizada na execução do modelo define um menor consumo de energia, importante característica desejada no projeto. O tamanho do arquivo armazenado na memória secundária também é vital para estabelecer um menor tempo de subir modelo para RAM, obtendo um melhor tempo de resposta da aplicação. Com os resultados em mãos, podemos estabelecer uma comparação com os modelos já existentes na literatura a fim de validar a eficácia do projeto ou levar possíveis correções às fases anteriores para melhorias.

VIII. EXPERIMENTOS

Nesse tópico serão apresentados os datasets utilizados para validação, os detalhes de implementação e discussões dos experimentos realizados.

A. Datasets e Pré-processamento

Foram utilizados 5 datasets: MHEALTH [1]; PAMAP2 [16]; WHARF [2]; UTD-1 [3]; UTD-2 [3]. A tabela IV mostra os detalhes em relação a frequência, número de atividades,

número de amostras e se o dataset é balanceado no que diz respeito a quantidade de amostras de cada atividade. Mais detalhes são apresentados em [9].

Os datasets foram utilizados com as mesmas regras de validação exposta por Jordão et al. [9]. Nesse caso, os dados de sensores coletados foram processados de maneira a separá-los em intervalos de 5 segundos. O que significa que as medições de predições dessa seção estão se referenciando a uma janela de 5 segundos.

B. Implementação e Hardware

Os códigos foram implementados em keras. A rede é composta com um ensemble de redes com a seguinte configuração: Conv(32,2x2), Maxpool(2x1), Conv(16,2x2), Maxpool(2x1), Dense(200), dropout(0.6), Dense(200), dropout(0.6), Concatenate(), Dense(200), dropout(0.6), Dense(200), dropout(0.6), softmax. Os hardware utilizados para comparar o desempenho do modelo em dispositivo móvel foram raspberry 3 e raspberry 4. Os detalhes desses hardwares podem ser observados na tabela III.

C. Resultados e Discussões

Os resultados são apresentados na tabela VII, em acurácia e f1. O valor de f1 foi utilizado devido ao desbalanceamento de alguns datasets, que comprometem a comparação utilizando a acurácia. Com o intervalo de confiança de 95%.

Essa tabela VII apresenta o resultado da avaliação de comparação entre os métodos da literatura, o nosso método e variações da nossa proposta utilizando o acelerômetro do dispositivo do braço. Utilizamos somente o acelerômetro do braço devido a utilização desses dados pela literatura. Nosso método com as 3 representações se equipara ou é melhor em quase todos os datasets. Ficando atrás somente no dataset UTD-1.

A tabela V e VI apresentam os dados da execuções e comparação das informações de utilização de recurso computacional. Comparação entre os métodos apresenta que o pré-processamento que necessita gerar o plote de recorrência tem muito custo computacional e demora um tempo considerável.

Dado os resultados de acurácia e de recurso computacional, pode ser deixado de fora um pré-processamento ou outro.

IX. VALIDAÇÃO E VERIFICAÇÃO

O processo de teste abrange as etapas de teste de desenvolvimento, teste de sistema e teste de aceitação, desde o início do processo até a garantia de cumprimento dos requisitos. A verificação e validação foram executadas durante e depois do desenvolvimento seguindo-se o Modelo V, desde a especificação dos requisitos, revisões de projeto e inspeções e testes de código. A verificação foi feita utilizando-se da comparação com métodos do estado da arte, permitindo as mesmas condições, mesmo conjunto de dados e estrutura de validação. A validação baseou-se no método Leave-One-Out, onde os dados de cada pessoa foram utilizados como teste uma vez e o resultado final é a média das execuções. Foram utilizadas as métricas de acurácia, revocação, f1 e intervalo

de confiança. A comparação de desempenho em hardware foi feita em um dispositivo de baixo custo, comparável com o sistema que os métodos serão aplicados seguindo as métricas de porcentagem do processamento utilizado, porcentagem de memória RAM utilizada, tempo de processamento do conjunto de dados, tamanho de arquivo em memória secundária e tempo de subir modelo.

X. ADAPTAÇÕES DO SISTEMA

Durante a execução do projeto foram feitas mudanças. Inicialmente, foi pensado em usar muitos sensores, porém devido ao custo computacional de processar vários sensores, foi avaliado que poderia ser usado somente um sensor. Dessa forma, foi selecionado o acelerômetro por ser um sensor mais usado no geral, assim contendo mais informações.

Outra mudança é que a aplicação pode ser utilizada com permutação entre duas representações.

TABLE III
ESPECIFICAÇÕES DOS HARDWARES USADOS

Modelo	SoC	CPU	GPU	RAM
Raspberry Pi 3	Broadcom BCM2837	4x ARM Cortex-A53, 1.2GHz	Broadcom VideoCore IV	1GB LPDDR2 (900 MHz)
Raspberry Pi 4	Broadcom BCM2711	Quad core Cortex-A72 (ARM v8), 1.5 GHz	Broadcom VideoCore IV	2GB LPDDR4 -3200 SDRAM

TABLE IV

CARACTERÍSTICAS DOS DATASETS USADOS NOS EXPERIMENTOS DESSE TRABALHO. ESSA INFORMAÇÃO ESTÁ DISPONÍVEL EM JORDÃO ET AL. [9].

DATASET	FREQ (Hz)	#ATIVIDADES	#AMOSTRAS	BALANCEADO
MHEALTH[1]	50	12	2555	True
PAMAP2 [16]	100	10	7522	False
WHARF [2]	32	12	3871	False
UTD-1 [3]	50	21	3771	True
UTD-2 [3]	50	5	1137	True

possível mudança na implementação do projeto é propor uma forma mais rápida de se fazer o pré-processamento de dados.

XI. CUMPRIMENTO DOS REQUISITOS

O modelo construído conseguiu atender ao requisitos requeridos na implementação do projeto, são eles:

O modelo foi desenvolvido em Python e com o uso da biblioteca Keras (RNF007 e RNF008), para a validação e verificação do modelo utilizou-se dos raspberrys Pi 3 e 4. E, como visto pelas tabelas V e VI, o uso de memória RAM do modelo não passou de 150 MB e o uso de CPU ficou abaixo de 30% (RNF002,RNF003,RNF004,RNF005,RNF006). Também, como calculado, o tempo de predição médio do modelo para uma faixa de tempo de 5s é menor que 0.15s (RF002). Além disso, por estar sendo feito todo o processamento pelo dispositivo se cumpre o RF001. O modelo foi treinado em datasets com diversas atividades e para mais de um dataset como visto na tabela IV (RF003,RF006), tendo como saída a classificação da atividade (RF005). De forma a diminuir a complexidade do modelo os dados de entrada foram pré-processados antes de alimentarem a rede neural (RF007), o tempo de pré-processamento pode ser visto na tabela V e VI.

Por fim, em relação ao desempenho do modelo, o mesmo conseguiu se equiparar ao estado da arte e até mesmo ser melhor em alguns datasets como o MHEALTH e o PAMAP2 como mostrado pela tabela VII (RNF001 e RNF009).

XII. CONCLUSÃO

A partir do desenvolvido do trabalho foi possível trabalhar com diversos conceitos da Engenharia de Sistemas, como a análise de viabilidade do projeto e análise dos requisitos, modelagem e desenvolvimento do projeto, validação e verificação do cumprimento dos requisitos. Além disso, a dinâmica de stakeholders conseguiu deixar todas as etapas do projeto bem mais de acordo com a realidade.

A respeito dos resultados do projeto, acredita-se que foram bastante satisfatórios para o que foi proposto, conseguindo resultados equivalentes ao estado da arte e avaliando o modelo criado em hardware para simular como seria o desempenho do mesmo no mundo real, o que não é normalmente feito na literatura. Apesar disso, um ponto a se melhorar do projeto é o pré-processamento dos dados, que representa a maior parte do tempo de execução do algoritmo, o que pode torná-lo inviável em algumas situações reais. Dessa forma, uma

TABLE V

TABELA DE RECURSOS COMPUTACIONAIS, APRESENTA OS PARÂMETROS, TEMPO DE EXECUÇÃO, MEMÓRIA E USO DA CPU PARA CADA MÉTODO USANDO UM RASPBERRY3

MÉTODO	Parâmetros	Pré-processamento (s)	Predição (s)	Memória (MB)	CPU (%)
Sena et. al.	6,021,488	0.0004	0.4471	176.79565	26.8
Lu et. al.	59,048	16.0537	0.6968	136.76953	26.0
Rec. plot mod + spectro.	-	16.5688	0.2561	129.69165	24.8
Time + Rec. plot mod.	-	16.2951	0.1904	137.12179	26.1
Time + spectro.	-	0.0067	0.1531	135.1673	27.2
Ensemble 3 representations	4,561,456	16.2249	0.3151	140.04224	25.0

TABLE VI

TABELA DE RECURSOS COMPUTACIONAIS, APRESENTA OS PARÂMETROS, TEMPO DE EXECUÇÃO, MEMÓRIA E USO DA CPU PARA CADA MÉTODO USANDO UM RASPBERRY4

MÉTODO	Parâmetros	Pré-processamento (s)	Predição (s)	Memória (MB)	CPU (%)
Sena et. al.	6,021,488	0.0002	0.2163	176.72601	26.9
Lu et. al.	59,048	7.0704	0.3819	138.89945	26.1
Rec. plot mod + spectro.	-	6.8344	0.0892	146.06745	26.3
Time + Rec. plot mod.	-	6.9578	0.0887	149.38112	27.8
Time + spectro.	-	0.0032	0.0731	134.54950	28.9
Ensemble 3 representations	4,561,456	7.0412	0.1498	150.47884	25.5

TABLE VII

TABLE OF ACCURACY AND F1 SCORE OF THE METHODS PERFORMED FOR EACH DATASET

METHOD	MHEALTH		PAMAP2		UTD-1		UTD-2		WHARF	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Sena et.al. [19]	80.13±0.48	77.55±0.56	70.94±1.83	61.47±1.98	48.13±0.49	46.67±0.76	71.29±0.46	70.80±0.46	62.30±3.03	37.83±1.88
Lu et. al. [11]	73.02±0.76	69.61±0.96	52.55±0.98	46.92±1.21	36.32±0.40	34.83±0.50	76.56±0.26	76.10±0.27	62.73±0.73	35.40±0.25
Time repre.	75.22±0.56	72.09±0.36	62.29±2.24	49.17±1.56	46.32±0.44	45.13±0.50	73.59±0.57	73.45±0.60	50.96±3.55	27.40±4.02
Spsectro. repre.	68.86±1.36	63.54±2.15	64.87±1.53	53.04±2.42	26.41±0.06	25.19±0.08	56.70±0.25	54.28±0.24	48.63±0.91	26.18±0.52
Recurrence plot mod.	73.41±0.58	70.03±0.74	51.19±1.73	44.38±1.17	40.93±0.46	39.25±0.55	74.63±0.77	74.67±0.82	64.50±1.16	37.52±0.93
Time + spectro.	80.13±2.04	77.57±2.60	75.46±2.58	63.79±1.84	42.43±0.30	40.80±0.32	74.27±0.51	73.66±0.42	66.01±2.20	43.26±1.20
Time + Rec. plot mod.	81.65±0.87	77.50±0.85	63.66±1.44	53.54±0.90	38.51±0.75	36.95±0.70	75.95±1.13	75.92±1.13	61.78±1.45	38.81±0.88
Rec. plot mod + spectro.	73.83±0.45	72.12±0.55	61.61±0.89	54.67±0.56	35.64±0.56	34.62±0.55	73.29±0.86	72.88±0.96	62.89±0.88	42.89±2.02
Ensemble 3 representations	83.46±0.97	81.42±1.18	73.84±1.85	64.38±1.38	41.77±0.54	40.25±0.62	77.29±0.57	76.84±0.55	65.81±2.44	44.12±2.08

REFERÊNCIAS

- [1] Oresti Banos et al. “Design, implementation and validation of a novel open framework for agile development of mobile health applications”. Em: *BioMedical Engineering Online* 14.2 (2015), pp. 1–20. ISSN: 1475925X. DOI: 10.1186/1475-925X-14-S2-S6.
- [2] Barbara Bruno, Fulvio Mastrogiovanni e Antonio Sgorbissa. “A public domain dataset for ADL recognition using wrist-placed accelerometers”. Em: *IEEE RO-MAN 2014 - 23rd IEEE International Symposium on Robot and Human Interactive Communication: Human-Robot Co-Existence: Adaptive Interfaces and Systems for Daily Life, Therapy, Assistance and Socially Engaging Interactions* (2014), pp. 738–743. DOI: 10.1109/ROMAN.2014.6926341.
- [3] Chen Chen, Roozbeh Jafari e Nasser Kehtarnavaz. “UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor”. Em: *Proceedings - International Conference on Image Processing, ICIP 2015-December* (2015), pp. 168–172. ISSN: 15224880. DOI: 10.1109/ICIP.2015.7350781.
- [4] Kaixuan Chen et al. “Deep Learning for Sensor-based Human Activity Recognition: Overview, Challenges and Opportunities”. Em: 37.4 (2020). arXiv: 2001.07416. URL: <http://arxiv.org/abs/2001.07416>.
- [5] Jungong Han et al. “Enhanced computer vision with microsoft kinect sensor: A review”. Em: *IEEE transactions on cybernetics* 43.5 (2013), pp. 1318–1334.
- [6] Zhiwen Xiao Haoxi Zhang. “A Novel IoT-Perceptive Human ActivityRecognition (HAR) Approach Using Multi-Head Convolutional Attention”. Em: *IEEE Internet of Things Journal* (2019), pp. 1–9. DOI: 10.1109/jiot.2019.2949715.
- [7] Wenchao Jiang e Zhaozheng Yin. “Human activity recognition using wearable sensors by deep convolutional neural networks”. Em: *MM 2015 - Proceedings of the 2015 ACM Multimedia Conference* (2015), pp. 1307–1310. DOI: 10.1145/2733373.2806333.

- [8] Charmi Jobanputra, Jatna Bavishi e Nishant Doshi. “Human activity recognition: A survey”. Em: *Procedia Computer Science* 155.2018 (2019), pp. 698–703. ISSN: 18770509. DOI: 10.1016/j.procs.2019.08.100. URL: <https://doi.org/10.1016/j.procs.2019.08.100>.
- [9] Artur Jordao et al. “Human Activity Recognition Based on Wearable Sensor Data: A Standardization of the State-of-the-Art”. Em: (2018), pp. 1–12. arXiv: 1806.05226. URL: <http://arxiv.org/abs/1806.05226>.
- [10] Dahai Liu. *Systems Engineering: Design Principles and Models*. CRC Press, 2018. ISBN: 978-1466506831.
- [11] Jianjie Lu e Kai Yu Tong. “Robust Single Accelerometer-Based Activity Recognition Using Modified Recurrence Plot”. Em: *IEEE Sensors Journal* 19.15 (2019), pp. 6317–6324. ISSN: 15581748. DOI: 10.1109/JSEN.2019.2911204.
- [12] James Manyika et al. “THE INTERNET OF THINGS: MAPPING THE VALUE BEYOND THE HYPE”. Em: (). URL: <https://www.mckinsey.com/~media/McKinsey/Industries/Technology%20Media%20and%20Telecommunications/High%20Tech/Our%20Insights/The%20Internet%20of%20Things%20The%20value%20of%20digitizing%20the%20physical%20world/The-Internet-of-things-Mapping-the-value-beyond-the-hype.pdf>.
- [13] Carlos Mateo e Juan Antonio Talavera. “Short-Time Fourier Transform with the Window Size Fixed in the Frequency Domain (STFT-FD): Implementation”. Em: *SoftwareX* 8 (2018), pp. 5–8. ISSN: 23527110. DOI: 10.1016/j.softx.2017.11.005. URL: <https://doi.org/10.1016/j.softx.2017.11.005>.
- [14] Alan V Oppenheim. *Discrete-time signal processing*. Pearson Education India, 1999.
- [15] Preksha Pareek e Ankit Thakkar. “A survey on video-based Human Action Recognition: recent updates, datasets, challenges, and applications”. Em: *Artificial Intelligence Review* (2020), pp. 1–64.
- [16] Attila Reiss e Didier Stricker. “Introducing a new benchmarked dataset for activity monitoring”. Em: *Proceedings - International Symposium on Wearable Computers, ISWC* (2012), pp. 108–109. ISSN: 15504816. DOI: 10.1109/ISWC.2012.13.
- [17] Fernando Moya Rueda et al. “Convolutional neural networks for human activity recognition using body-worn sensors”. Em: *Informatics* 5.2 (2018), pp. 1–17. ISSN: 22279709. DOI: 10.3390/informatics5020026.
- [18] Olli Särkkä et al. “A multi-position calibration method for consumer-grade accelerometers, gyroscopes, and magnetometers to field conditions”. Em: *IEEE Sensors Journal* 17.11 (2017), pp. 3470–3481.
- [19] Jessica Sena, Jesimon Barreto Santos e William Robson Schwartz. “Multiscale DCNN ensemble applied to human activity recognition based on wearable sensors”. Em: *European Signal Processing Conference 2018-Septe* (2018), pp. 1202–1206. ISSN: 22195491. DOI: 10.23919/EUSIPCO.2018.8553354.
- [20] Jessica Sena et al. “Human activity recognition based on smartphone and wearable sensors using multiscale DCNN ensemble”. Em: *Neurocomputing* xxxx (2020). ISSN: 18728286. DOI: 10.1016/j.neucom.2020.04.151. URL: <https://doi.org/10.1016/j.neucom.2020.04.151>.
- [21] Ji Wang et al. “Deep learning towards mobile applications”. Em: *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*. IEEE. 2018, pp. 1385–1393.
- [22] Lei Wang, Du Q Huynh e Piotr Koniusz. “A comparative review of recent kinect-based action recognition algorithms”. Em: *IEEE Transactions on Image Processing* 29 (2019), pp. 15–28.
- [23] Bendong Zhao et al. “Convolutional neural networks for time series classification”. Em: *Journal of Systems Engineering and Electronics* 28.1 (2017), pp. 162–169. ISSN: 16711793. DOI: 10.21629/JSEE.2017.01.18. arXiv: 1603.06995.