

**REDE DOCTUM DE ENSINO
FACULDADE INTEGRADAS DE CARATINGA
CURSO SUPERIOR DE ENGENHARIA CIVIL**

**PROTÓTIPO DE SISTEMA DE ANÁLISE AUTOMATIZADA
DE ATIVIDADES PARA A CONSTRUÇÃO CIVIL**

**IZAQUE JOSÉ DA SILVA JUNIOR
LEANDRO NUNES BATISTA**

Trabalho de Conclusão de Curso

Caratinga/MG

2017

**IZAQUE JOSÉ DA SILVA JUNIOR
LEANDRO NUNES BATISTA**

**PROTÓTIPO DE SISTEMA DE ANÁLISE AUTOMATIZADA
DE ATIVIDADES PARA A CONSTRUÇÃO CIVIL**

Trabalho de Conclusão de Curso apresentado à Banca Examinadora do Curso Superior de Engenharia Civil das Faculdades Integradas de Caratinga da DOCTUM Caratinga como requisito parcial para obtenção do Grau de Bacharel em Engenharia Civil.

Professor Orientador: Dr. Douglas Borges Domingos.

Caratinga/MG

2017

TERMO DE APROVAÇÃO

TERMO DE APROVAÇÃO

O Trabalho de Conclusão de Curso intitulado: **PROTÓTIPO DE SISTEMA DE ANÁLISE AUTOMATIZADA DE ATIVIDADES PARA A CONSTRUÇÃO CIVIL**, elaborado pelos alunos **IZAQUE JOSÉ DA SILVA JUNIOR** e **LEANDRO NUNES BATISTA** foi aprovado por todos os membros da Banca Examinadora e aceita pelo curso de Engenharia Civil das **FACULDADES DOCTUM DE CARATINGA**, como requisito parcial da obtenção do título de

BACHAREL EM ENGENHARIA CIVIL.

Caratinga 10 de julho 2017



DOUGLAS BORGES DOMINGOS



JOSÉ SALVADOR ALVES



CAMILA ALVES DA SILVA

*À Deus por sua infinita misericórdia e graça,
à família pelo amparo e cuidado, e aos amigos pela presença e companheirismo.*

Izaque José da Silva Junior

*A todos amigos, professores e colegas que direta ou indiretamente puderam
contribuir com apoio e confiança durante esta jornada. Dedico especialmente à minha família,
pela paciência, confiança e incentivo, sem a qual eu jamais poderia ter chegado até aqui.*

Leandro Nunes Batista

AGRADECIMENTOS

A Deus, "em quem estão escondidos todos os tesouros da sabedoria e do conhecimento"(Cl 2.3), o meu muitíssimo obrigado, o meu limitado vocabulário não consegue expressar o poderio e soberania de um Deus tão grande. Por sempre estar comigo, ser o meu socorro, refúgio, amigo, por ter me salvado, eu te agradeço Deus. Pelas oportunidades e bênçãos concedidas, pelo cuidado para com a minha vida, mesmo em terras longínquas, sempre se fez presente. Confiar em Ti, nos faz ter a certeza de que por mais difícil seja a caminhada, no final, vai valer a pena.

Ao meu companheiro de pesquisa e amigo Leandro, por sua enorme contribuição e paciência comigo, no desenvolvimento deste trabalho. Muito obrigado pela confiança e ajuda.

Agradeço também aos amigos que fiz no meu período de estágio na Prefeitura Municipal de Caratinga, por todo o conhecimento e coleguismo a mim dispensados. Em especial ao Engenheiro Wladimir, por seu incentivo e experiência transmitidos; aos amigos Rodrigo, Sara, Juan, Leonardo e Carlos, pelo companheirismo e aprendizado; e a todos da Assessoria de Projetos, por terem sempre me apoiado e pela recepção tão agradável.

A todos os professores e colegas da faculdade, que fizeram parte dessa árdua caminhada e pela contribuição no meu crescimento e desenvolvimento pessoal e profissional. Em especial ao Marcus Vinícius, José Paulo, Anacyr, Fernando, dentre outros, pelo conhecimento e amizade comigo compartilhados.

To the friends I have made abroad, who have always given me encouragement and support. To all the friends I made at Christian Student Fellowship, Ashland Baptist Church, Coral Gables Baptist Church, University of Kentucky and University of Miami, friends from all over the world who were always willing to help me. I am forever grateful for the things you have taught me and the friendship you have given me.

Aos irmãos e amigos da Igreja Assembleia de Deus do Bairro Limoeiro, por seu apoio e orações. A Cálita Cristina, por toda sua contribuição e vontade, incentivo para que eu pudesse continuar seguindo em frente.

À minha família, por serem meus maiores incentivadores, apoiadores, mentores e amigos. Em especial aos meus pais Izaque José da Silva e Flávia Medeiros da Silveira e Silva, pela confiança, amor, exemplo, e por tudo que representam na minha vida. Aos meus amigos e amados irmãos Samuel Igor e Davi Ícaro, por serem tão pacientes e bondosos comigo. A minha família é o meu bem mais precioso e também a minha maior motivação.

Izaque José da Silva Júnior

Agradeço a Deus por toda a sua graça e sabedoria e, principalmente, por ter proporcionado a luz que tem me guiado a cada passo dado durante esta jornada.

Agradeço a cada um dos amigos que fiz durante todos os anos deste curso, por proporcionarem momentos que deverão ser guardados na memória pelo resto da minha vida. Fossem esses momentos apenas alguns minutos de descontração, ou experiências únicas e marcantes em nossas vidas.

Ao meu amigo e companheiro de pesquisa, Izaque, por ter tido paciência nos momentos de cobrança e estresse ao longo do desenvolvimento deste trabalho.

Agradeço a todos os professores pela paciência e, principalmente, pela generosidade daqueles que em vários momentos forneceram apoio e conhecimento muito além do que apenas poderia ser previsto pela função.

Agradeço especialmente à minha família, pelo apoio e amor incondicional sempre presente em minha vida. Agradeço por terem depositado em mim, em todos os momentos, a mais plena confiança, mesmo quando eu estava cheio de insegurança e incerteza. Agradeço, sobretudo, por permitirem que eu aprendesse lições que carrego para a vida e que nenhum livro ou curso seria capaz de ensinar.

Leandro Nunes Batista

*“A tarefa não é tanto ver aquilo que ninguém viu,
mas pensar o que ninguém ainda pensou sobre aquilo que todo mundo vê.”*
(ARTHUR SCHOPENHAUER)

BATISTA, Leandro Nunes; DA SILVA JUNIOR, Izaque José. **PROTÓTIPO DE SISTEMA DE ANÁLISE AUTOMATIZADA DE ATIVIDADES PARA A CONSTRUÇÃO CIVIL**. Caratinga, 2017. 51f. Trabalho de Conclusão de Curso Superior de Engenharia Civil - Curso de Engenharia Civil. Faculdades Integradas de Caratinga, Rede DOCTUM, Caratinga, 2017.

RESUMO

No ramo da construção civil, a mão-de-obra se destaca por ser um insumo de representatividade considerável no custo final dos empreendimentos. A demanda por esse insumo é grande e necessária, portanto, os serviços prestados devem ser produtivos e de boa qualidade. Ciente da escassez de técnicas acessíveis para controle e monitoramento das atividades realizadas no canteiro de obras, este trabalho objetiva analisar a viabilidade do desenvolvimento e implementação de um protótipo capaz de reconhecer e classificar as atividades que estão sendo executadas. Utilizando-se das técnicas de *work sampling*, sistema de posicionamento interno, e reconhecimento automático de atividades humanas, testes foram realizados para a avaliação do funcionamento do sistema. Os testes foram realizados em três etapas, sendo estas: a identificação do local onde estavam sendo desenvolvidas as atividades; a amostragem e identificação das atividades propriamente ditas; e a análise da produtividade. Os resultados obtidos se mostraram interessantes, mas com restrições relativas ao equipamento utilizado -um *smartphone* neste caso-, e condições do ambiente, uma vez que o trabalho foi desenvolvido em um ambiente controlado e a sua implementação em uma situação real, exigiria um aprimoramento do sistema para que houvesse maior precisão. O sistema se mostrou acessível, passível de aprimoramento, e funcional, sendo capaz de analisar as atividades, contribuindo para o gerenciamento de projetos na construção civil.

Palavras-chave: Amostragem de trabalho. Análise de Atividades. Automação. Construção Civil. Mão-de-obra. Produtividade. Sistema de Posicionamento.

BATISTA, Leandro Nunes; DA SILVA JUNIOR, Izaque José. **AUTOMATED ACTIVITY ANALYSIS PROTOTYPE FOR CIVIL CONSTRUCTION**. Caratinga, 2017. 51f. Trabalho de Conclusão de Curso Superior de Engenharia Civil - Curso de Engenharia Civil. Faculdades Integradas de Caratinga, Rede DOCTUM, Caratinga, 2017.

ABSTRACT

In the construction sector, the workforce stands out as an input of considerable representativeness in the final cost of the projects. The demand for this input is great and necessary, therefore, the services provided must be productive and of good quality. Aware of the scarcity of accessible techniques to control and monitor the activities carried out at the construction site, this work aims to analyze the feasibility of the development and implementation of a prototype capable of recognizing and classifying the activities that are being carried out. Using the techniques of work sampling, internal positioning system, and automatic recognition of human activities, tests were carried out to evaluate the functioning of the system. The tests were performed in three stages, these being: the identification of the place where the activities were being developed; the sampling and identification of the activities themselves; and productivity analysis. The results obtained were interesting, but with restrictions related to the equipment used - a smartphone in this case - and environmental conditions, since the work was developed in a controlled environment and its implementation in a real situation, would require an improvement of the System for greater precision. The system was accessible, susceptible of improvement, and functional, being able to analyze the activities, contributing to the management of projects in civil construction.

Key-words: Activities Analysis. Automation. Civil Construction. Positioning System. Productivity. Workforce. Work Sampling.

LISTA DE ILUSTRAÇÕES

Figura 1	– Ilustração de modelo de IPS utilizando a técnica de <i>fingerprinting</i>	27
Figura 2	– Demonstração do posicionamento do <i>smartphone</i> no corpo do usuário. . .	28
Figura 3	– Demonstração do processo de classificação dos dados dos sensores.	29
Figura 4	– Local designado para armazenamento de material.	35
Figura 5	– Local designado para descanso.	35
Figura 6	– Local designado para execução do trabalho.	36
Figura 7	– Execução do Serviço de Alvenaria.	36
Figura 8	– Matriz de Confusão Normalizada.	38
Figura 9	– Fluxograma do Sistema de Classificação Final.	40

LISTA DE TABELAS

Tabela 1 – Sinais no domínio de tempo e frequência obtidos à partir dos sensores de smartphone	31
Tabela 2 – Avaliação das classificações iniciais para determinação das classificações finais	33
Tabela 3 – Comparação das Classificações Finais Manuais com Base em Vídeo e a Classificação Automática Realizada Pelo Sistema.	41

LISTA DE ABREVIATURAS E SIGLAS

IPS	<i>Internal Positioning System</i> - Sistema de Posicionamento Interno
GPS	<i>Global Positioning System</i> - Sistema de Posicionamento Global
RRSI	<i>Received Signal Strength Indication</i> - Indicador de Intensidade de Sinal Recebido
MFP	<i>Multi Factor Productivity</i> - Produtividade Multifatorial
SFP	<i>Single Factor Productivity</i> - Produtividade Uni-fatorial
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
TCPO	Tabela de Composições de Preços para Orçamentos
PMM	<i>Productivity Measuring Methods</i> - Métodos de Mensuração de Produtividade
PIM	<i>Productivity Improving Methods</i> - Métodos de Aprimoramento de Produtividade
MAP	<i>Maximum a Posteriori</i> - Máximo a Posteriori
DTC	<i>Decision Tree Classifier</i> - Classificador de Árvore de Decisão
CART	<i>Classification and Regression Tree</i> - Árvore de Regressão e Classificação
CSV	<i>Comma Separated Values</i> - Valores Separados por Vírgula

LISTA DE SÍMBOLOS

Π	Letra Grega Pi
ω	Letra Grega Ômega
Σ	Letra Grega Sigma

SUMÁRIO

1 INTRODUÇÃO	15
1.1 OBJETIVO GERAL	16
1.2 OBJETIVOS ESPECÍFICOS	16
1.3 JUSTIFICATIVA	16
1.4 DELIMITAÇÃO	17
1.5 ESTRUTURA	17
2 REVISÃO DA LITERATURA	18
2.1 TÉCNICAS E TECNOLOGIAS PARA MENSURAÇÃO DA PRODUTIVIDADE	18
2.1.1 Escolha de parâmetros para mensuração de produtividade	19
2.1.2 PMM e PIM	20
2.1.3 Work sampling	20
3 PROCEDIMENTOS METODOLÓGICOS	23
3.1 DESENVOLVIMENTO DO SISTEMA PROTÓTIPO	23
3.1.1 Sistema de Posicionamento Interno com base em redes Wi-Fi	25
3.1.2 Fingerprinting	26
3.1.2.1 Classificador de Naive Bayes	27
3.2 SISTEMA DE DETECÇÃO DE MOVIMENTOS COM BASE EM DADOS DE ACELERÔMETRO E GIROSCÓPIO	28
3.2.1 Coleta de Dados	28
3.2.2 Classificação Manual	29
3.2.3 Análise e limpeza de dados	30
3.2.4 Classificador de Árvore de Decisão	31
3.2.5 Treinamento do Classificador	32
3.3 INTERSEÇÃO DOS DADOS DE MOVIMENTAÇÃO E LOCALIZAÇÃO	32
3.4 APRESENTAÇÃO DAS BIBLIOTECAS UTILIZADAS	33
3.4.1 Scikit-learn	33
3.4.2 Pandas	33
3.4.3 NumPy	33
3.4.4 SciPy	34
3.4.5 Matplotlib	34
3.5 APRESENTAÇÃO DO EXPERIMENTO PROPOSTO	34
4 RESULTADOS E DISCUSSÕES	37
4.1 PRECISÃO DO CLASSIFICADOR	37
4.2 PRECISÃO DO SISTEMA DE LOCALIZAÇÃO INTERNA	39
4.3 PRECISÃO DAS CLASSIFICAÇÕES FINAIS	39

5 CONCLUSÃO	42
6 REFERÊNCIAS BIBLIOGRÁFICAS	44
APÊNDICE A Programa de própria autoria para processamento dos sinais recebidos pelo <i>smartphone</i> e reconhecimento de atividades	46

1 INTRODUÇÃO

Quando se fala de execução de obras na construção civil, a mão-de-obra tem lugar de destaque, a sua representatividade no custo final dos empreendimentos é considerável, o que demanda que a tal seja de boa qualidade e eficiente. Diversos fatores influenciam na qualidade da mão-de-obra – salários em dia, saúde ocupacional, motivação no trabalho são exemplos –, e na sua respectiva produtividade, sendo que a sua mensuração é de fundamental importância [1].

No Brasil, o aparato tecnológico usado na construção civil, nas últimas décadas e de um modo geral, era bastante limitado, sendo que grande parte do processo construtivo se dava de modo manual e empírico, isso aconteceu devido ao baixo investimento de tempo e dinheiro no ramo tecnológico aplicado à construção. Com o crescimento econômico do país na última década, nota-se então maiores investimentos em capital físico, tendo como consequência, uma melhora na qualidade da mão-de-obra e sua produtividade [2].

Estes avanços refletem também nos investimentos em novas tecnologias e numa maior receptividade à inovação. A aplicação de novas tecnologias em setores antes dominados por processos tradicionais agora torna-se uma estratégia competitiva altamente eficiente [3].

Outro fator importante a ser considerado no investimento e adoção de novas tecnologias, é a diminuição de seu custo e complexidade de implementação. Sistemas que antes exigiam uso de equipamentos especializados, hoje podem ser amplamente encontrados em dispositivos portáteis, como os *smartphones*. Um exemplo de tecnologia que tem se aprimorado e tornado mais popular é a de sistemas de posicionamento interno. Um Sistema de Posicionamento Interno (*Internal Positioning System*, IPS) é um sistema que utiliza ondas de rádio, campos magnéticos ou qualquer outra informação de sensores para localizar pessoas ou objetos dentro de um ambiente fechado [4].

Um IPS, em contraste a um Sistema de Posicionamento Global (*Global Positioning System*, GPS), não depende de satélites e, portanto, torna possível a localização em ambientes fechados ou subterrâneos, onde a intensidade dos sinais de satélite é baixa ou nula.

Atualmente os IPSs utilizam uma variedade de técnicas para possibilitar o fornecimento de informações de localização com um nível mínimo de precisão. A mais comum delas é a trilateração a partir do Indicador de Intensidade de Sinal Recebido (*Received Signal Strength Indication*, RSSI). No entanto, outras técnicas mais recentes também têm sido utilizadas com a capacidade de aumentar significativamente a precisão dos resultados [3].

A aplicação deste tipo de sistema na construção civil oferece diversos benefícios ao setor, entretanto sua utilização ainda é restrita a um número limitado de soluções comerciais disponíveis atualmente [5].

A técnica de análise de atividades tem como objetivo principal, o reconhecimento e análise das atividades sendo desenvolvidas a partir de um vídeo contínuo, automaticamente. De uma forma simplificada, uma vez que as atividades são definidas previamente, ao serem executadas por um trabalhador, o sistema precisa detectar e classificar o vídeo em cada uma das categorias definidas anteriormente [6].

O presente trabalho aborda as possíveis aplicações de tal sistema dentro da construção civil e como os dados de localização e movimentação obtidos dentro do canteiro de obras podem ser utilizados para auxiliar no monitoramento de produtividade de trabalhadores dentro do canteiro de obras.

Além da análise da aplicabilidade de Sistemas de Posicionamento Interno e análise de atividades no setor da construção civil, o presente trabalho apresenta o processo de elaboração de protótipo de um Sistema de Posicionamento Interno de baixo custo voltado para a construção civil.

1.1 OBJETIVO GERAL

Avaliar a implementação um sistema automatizado de análise de atividades na construção civil.

1.2 OBJETIVOS ESPECÍFICOS

- Desenvolver um protótipo de Sistema de Posicionamento Interno e análise automatizada de atividades para implementação em um canteiro de obras;
- Realizar testes com o protótipo do sistema;
- Analisar a viabilidade da utilização do sistema para monitoramento de execução de atividades com base nos testes realizados e na revisão bibliográfica de material científico disponível.

1.3 JUSTIFICATIVA

A utilização de dados para a tomada de decisões vem se tornando uma prática padrão dentro de diversos setores da indústria. A captura e análise de dados tem se tornado uma ferramenta de grande valor dentro das empresas, reduzindo custos, elevando níveis de eficiência e aumentando a produtividade.

A construção civil, no entanto, ainda se baseia fortemente em processos manuais e tende a depender de mão de obra qualificada para a execução adequada de tais processos. No entanto,

poucos dados são aferidos sobre a execução do trabalho dentro do canteiro de obras e quando estes dados são aferidos, o procedimento é realizado com pouco auxílio de tecnologia.

Capturar e analisar dados referentes a localização e movimentação de trabalhadores dentro de um canteiro de obras apresenta-se como uma interessante ferramenta capaz de fornecer informações importantes sobre produtividade, segurança e logística. Este trabalho propõe, portanto, utilizar um sistema de posicionamento interno para capturar dados de localização de trabalhadores de uma obra de construção civil e utilizar tais dados para determinar se um operário está executando trabalho, com base no sistema desenvolvido para detecção de seus movimentos e monitoramento de sua respectiva localização e da implementação da técnica de análise de atividade.

1.4 DELIMITAÇÃO

O projeto terá sua limitação definida na elaboração do protótipo do sistema de monitoramento e detecção de movimento e a avaliação da viabilidade da implantação e aplicação no ramo da construção civil.

1.5 ESTRUTURA

O primeiro capítulo deste trabalho apresenta de forma resumida, os desafios encontrados no mercado da construção civil quando se trata de monitoramento de atividades dentro do canteiro de obras, utilizando-se de aparatos tecnológicos. No mesmo capítulo são descritos os objetivos a serem alcançados, justificativa e delimitação deste projeto.

No segundo capítulo, técnicas e tecnologias para mensuração da produtividade são apresentadas, e conceitos considerados "chave" para o desenvolvimento do trabalho são descritos.

O terceiro capítulo aborda qual foi a metodologia que norteou este projeto, e a execução de seu desenvolvimento; bem como a introdução de conceitos mais aprofundados, essenciais para o bom entendimento do mesmo.

O quarto capítulo apresenta os resultados obtidos através dos testes realizados, contendo os detalhes da precisão das classificações efetuadas.

O quinto capítulo apresenta em suma, as considerações finais sobre o projeto, suas limitações, vantagens e desafios futuros, com a possibilidade de aprimoramento do sistema para uma contribuição efetiva na construção civil.

No Apêndice A, é possível encontrar os códigos, na linguagem Python, utilizados para o desenvolvimento do trabalho.

2 REVISÃO DA LITERATURA

O processo produtivo sempre fez parte da vida humana, desde os primórdios da humanidade a produção é notoriamente ativa, no sentido de que o homem primitivo produzia, a partir de rochas, pedaços de madeira e outros materiais, instrumentos que o auxiliariam na sua proteção e de sua família. Milhares de anos se passaram desde então, até o surgimento de métodos produtivos rápidos e eficientes.

Desse modo, surge portanto, o método de produção em massa. Tal método se tornou um processo econômico, pois diminuía consideravelmente a mão-de-obra, um dos principais recursos no processo produtivo, responsável por uma parcela significativa do custo final de produtos e serviços. A indústria seriada, em contraste a indústria da construção civil, vem estudando a aplicação do tema da produtividade nos métodos de produção em massa há mais de um século, como se observa através das obras de Taylor, escritas no final do século XIX, e Gilbreth, no início do século XX. Tais estudos foram capazes de aumentar a produtividade da indústria seriada em mais de 50 vezes [7].

A produtividade determina, portanto, se recursos são utilizados, ou não, de forma eficiente na produção de bens e serviços consumíveis. Um processo no qual os recursos são utilizados econômica e responsabilmente, de modo a resultar em produtos e serviços de qualidade, realizados em tempo hábil, é considerado produtivo [8].

Deste modo, com o desenvolvimento de novas tecnologias e o aumento da competitividade, é fundamental para as empresas que a sua produtividade seja eficiente e bem gerenciada. Com isto em mente, três procedimentos básicos para a gestão da produtividade devem ser considerados: (1) a mensuração da produtividade; (2) a determinação e análise dos obstáculos da produtividade; (3) a proposição e aplicação de soluções para a superação de tais obstáculos [9].

Na construção civil, a importância da produtividade é tema de estudo de diversos pesquisadores nacionais e internacionais. Tal importância se dá pelo peso da mão de obra no custo final de um empreendimento do setor. Este valor pode chegar a 50% do custo total, dependendo do grau de industrialização e tipo de obra [10]. Sua representatividade no custo final da execução de obras é, portanto, muito importante e demonstra a sua complexidade por se tratar de um serviço prestado por pessoas, com diferentes origens e formações, que possuem necessidades a serem supridas [11].

2.1 TÉCNICAS E TECNOLOGIAS PARA MENSURAÇÃO DA PRODUTIVIDADE

Dentre os diversos fatores capazes de afetar significativamente as chances de sucesso de um projeto, quatro aspectos são considerados, de comum acordo entre pesquisadores, como os

mais importantes: custo, qualidade, tempo e segurança [12].

Observa-se, no entanto, que a obtenção da qualidade, custo e tempo apropriados para um projeto dependem diretamente da obtenção de níveis de produtividade igualmente apropriados. A produtividade, portanto, é amplamente utilizada como importante indicador de performance de um projeto nas mais diversas operações envolvidas em todas as etapas da construção. Não há, no entanto, na indústria ou na área acadêmica, consenso quanto a uma definição ou método padrão para avaliação de níveis de produtividade [13].

Um método comum para a avaliação da produtividade envolve a determinação da razão entre a saída de um processo produtivo (seu resultado) e sua entrada (Trabalho, equipamentos e materiais), assim como demonstrado na equação 2.1

$$\text{Fator de Produtividade} = \frac{\text{Saída Física (Unidade)}}{\text{Trabalho(R\$)} + \text{Equipamento(R\$)} + \text{Materiais(R\$)}} \quad (2.1)$$

2.1.1 Escolha de parâmetros para mensuração de produtividade

A escolha dos parâmetros que irão compor a análise produtiva depende do escopo do processo produtivo e da disponibilidade de dados. Deste modo, no quesito escolha de parâmetros, é possível classificar as análises produtivas em dois grupos: o grupo de análises multifatoriais (MFP, ou *Multi Factor Productivity*) e o grupo de análises uni-fatoriais (SFP, ou *Single Factor Productivity*) [13]. As análises multifatoriais, por definição, levam em consideração mais de um parâmetro para se obter um valor capaz de indicar o nível de produtividade da atividade na qual a análise está envolvida. Os parâmetros utilizados em uma análise multifatorial podem envolver dados referentes a quantidades e/ou custos de materiais, equipamentos, energia e qualquer outro aspecto envolvido na obtenção do resultado final da atividade analisada. A equação 2.1 é um exemplo de análise de produtividade multifatorial.

Em uma análise de produtividade uni-fatorial, também por definição, apenas um parâmetro é levado em consideração para se determinar um fator capaz de indicar o nível de produtividade da atividade avaliada. As análises de produtividade uni-fatoriais também são conhecidas como análises parciais e representam uma forma mais econômica e eficiente de se mensurar a produtividade, uma vez que uma análise completa de todos os parâmetros envolvidos em qualquer processo produtivo eleva consideravelmente a complexidade desta análise.

A análise da produtividade da mão de obra é um exemplo de análise uni-fatorial que, de acordo com Organização para a Cooperação e Desenvolvimento Econômico (OCDE), é a mais utilizada no mundo [14]. A produtividade da mão de obra não leva em consideração, de modo explícito, o custo da mão de obra, assim como demonstra a equação 2.2

$$\text{Produtividade da Mão de obra} = \frac{\text{Horas Trabalhadas}}{\text{Unidade de Produção}} \quad (2.2)$$

No Brasil, a Tabela de Composições de Preços para Orçamentos (TCPO), da Editora PINI utiliza a equação 2.3 para referência de produtividade de diferentes tipos de serviços.

$$Produtividade = \frac{HomemHora}{Unidade\ de\ Produção\ (m^2, m^3, etc.)} \quad (2.3)$$

2.1.2 PMM e PIM

No quesito finalidade, os métodos de análise de produtividade podem ser divididos em dois grupos principais: Métodos de Mensuração de Produtividade (*Productivity Measuring Methods*, ou PMMs) e Métodos de Aprimoramento de Produtividade (*Productivity Improving Methods*, ou PIMs) [12].

No caso dos PMMs, os resultados das análises de produtividade são utilizados para referências internas ou externas, com fins de *benchmarking* [15], como é o caso da Tabela de Composições de Preços para Orçamentos (TCPO), da Editora PINI, na qual as taxas de produtividade são dadas na forma exemplificada pela equação 2.3. Tal tabela pode ser utilizada como referência para taxas de produtividade de diferentes tipos de serviços da construção civil.

Os PIMs, por outro lado, tem como intuito avaliar quão bem a utilização de equipamentos e mão de obra está sendo gerenciada [15]. Grande parte dos métodos que se encaixam na classificação de Métodos de Aprimoramento de Produtividade baseiam-se na utilização de estudos de tempo e movimento, feitos de modo manual com auxílio de cronômetros, questionários e entrevistas [12].

2.1.3 Work sampling

Um exemplo de PIM amplamente utilizado é a técnica conhecida como *work sampling* (amostragem do trabalho). Essa técnica se dá, através da utilização de uma grande quantidade de observações instantâneas, feitas durante determinado período de tempo, onde dados são coletados de trabalhadores, processos, máquinas etc. Ou seja, durante o período de amostragem, os dados são coletados no instante em que determinada tarefa está sendo executada, sendo o material de análise, a observação da atividade desenvolvida. A porcentagem total do registro de observações, compreende as porcentagens de tempo dispensados na execução de tarefas [16].

O objetivo principal do uso dessa técnica é avaliar como o tempo está sendo empregado pela mão-de-obra através da observação, classificação e coleta de dados percentuais de atividade de determinado projeto. Concomitantemente ao objetivo principal, a técnica é empregada a fim de que os problemas que interferem diretamente na fluidez do trabalho e que requerem atenção especial da parte gerencial, encontrados na execução das tarefas, sejam sanados de forma eficiente [16].

Através do *work sampling* é possível, estatisticamente, obter uma estimativa de duração de atividade, com a definição prévia de erro aceitável e grau de confiabilidade estatístico. Dessa forma, a amostragem é realizada em concordância com três regras fundamentais:

1. A instantaneidade das observações é de fundamental importância, devendo o observador analisar a atividade que está sendo desempenhada pelo trabalhador no primeiro momento que se olhar para o mesmo;
2. Não deve existir padrão de intervalo de observações a ser seguido, o que demanda que as observações sejam, essencialmente, aleatórias;
3. A fim de se atingir o erro aceitável e o grau de confiabilidade, previamente estabelecidos, é necessário que as observações sejam suficientes para representar o espaço amostral [17].

Para que a terceira regra seja atendida, é possível a realização do cálculo do número de observações necessárias para a representação da verdade, utilizando-se da Equação 2.4

$$N = \left(\frac{Z}{Er} \right)^2 * \left(\frac{1-p}{p} \right) \quad (2.4)$$

Onde:

N = número de observações necessárias

Z = coeficiente de distribuição normal

Er = erro relativo aceitável

p = proporção da atividade estudada no conjunto de todas as atividades [17].

Uma vez definido o número de observações a serem feitas, é importante a determinação de categorias que servirão de parâmetro para a análise em estudo:

1. Trabalho Direto – O tempo gasto diretamente na execução de um serviço;
2. Trabalho Indireto – todo o trabalho que contribui para a realização do Trabalho Direto, p. ex., transporte e manuseio de materiais e equipamentos no perímetro estabelecido como área de trabalho;
3. Trabalho Ineficiente – quando o tempo não é empregado para a realização do trabalho. Compreendendo o ato de ficar parado em estado ocioso, deslocamentos desnecessários no perímetro de trabalho etc [18].

Mesmo sendo amplamente utilizada, esta técnica, no entanto, apresenta diversas limitações e exige controle rigoroso em sua execução. Além disso, há o envolvimento de intenso trabalho manual que, conseqüentemente, resulta na ineficiência do provimento de resultados rápidos e precisos [12].

Estas mesmas limitações se aplicam a qualquer técnica não automatizada para aferimento de níveis de produtividade. Cabe, portanto, uma conclusão razoável de que a utilização de métodos capazes de automatizar a utilização destas técnicas é capaz de oferecer grandes benefícios para o setor da construção civil.

O sistema apresentado neste trabalho consiste numa implementação automatizada do método de análise de nível de atividade, que é considerada uma evolução da técnica de *work sampling* [12]. A análise de atividade é baseada na avaliação da taxa de tempo de trabalho direto, que descreve o tempo em que um trabalhador dedica diretamente a execução de uma determinada atividade [14]. Esta avaliação é demonstrada na Equação 2.5.

$$\text{Taxa de Tempo de Trabalho Direto} = \frac{\text{Tempo de Trabalho Direto}}{\text{Tempo Total de Trabalho}} \quad (2.5)$$

Não há, no entanto, evidência de uma total correlação entre a taxa de tempo de trabalho direto e a produtividade real de um trabalhador. Contudo, a utilização da taxa de tempo de trabalho direto é uma ferramenta que representa uma importante evolução quando comparada a técnicas tradicionais de amostragem de trabalho, sendo capaz de fornecer informações valiosas sobre a eficiência no gerenciamento e utilização do tempo dos trabalhadores em um canteiro de obras [12].

3 PROCEDIMENTOS METODOLÓGICOS

3.1 DESENVOLVIMENTO DO SISTEMA PROTÓTIPO

Para este trabalho um sistema protótipo foi desenvolvido integrando dois sistemas principais: um sistema de posicionamento interno baseado em redes Wi-Fi, e um sistema de reconhecimento automático de atividades humanas, baseado em informações de acelerômetro e giroscópio de um *smartphone* acoplado ao cinto de um usuário.

O sistema de posicionamento interno é constituído pelos seguintes componentes:

1. Um aplicativo de código livre baseado em um *framework*, ou biblioteca, de localização interna utilizando redes Wi-Fi desenvolvido para smartphones com sistema operacional Android e modificado para integração com o sistema de reconhecimento de atividades;
2. Um programa instalado em um servidor com sistema operacional Linux responsável pela comunicação com o aplicativo instalado no smartphone, bem como pelo gerenciamento do sistema de localização interna;

O sistema de reconhecimento de atividades humanas é, de modo semelhante ao anterior, constituído pelos seguintes componentes:

1. Um aplicativo instalado em um smartphone com sistema operacional Android responsável pela leitura e envio, para o servidor, dos dados de acelerômetro e giroscópio do smartphone;
2. Um programa instalado em um servidor com sistema operacional Linux constituído por um componente de treinamento, um componente de leitura e armazenamento dos dados recebidos do smartphone e um componente de processamento e classificação destes dados. Todos os componentes foram desenvolvidos na linguagem de programação Python em conjunto com bibliotecas de código livre disponíveis nesta mesma linguagem (Matplotlib, NumPy, scikit-learn, SciPy e Pandas). Uma breve apresentação das funções de cada uma destas bibliotecas é dada na Seção 3.4.

Um dos principais objetivos deste trabalho é demonstrar a aplicação das técnicas aqui mencionadas utilizando equipamento de baixo custo. A melhor alternativa encontrada atualmente reside na utilização de smartphones. Tais dispositivos, nos modelos mais modernos, possuem acelerômetros, giroscópio e módulo de Wi-Fi, sendo apenas estes os equipamentos necessários para a implementação das tecnologias aqui apresentadas.

Os acelerômetros são sensores que informam um valor real de aceleração ao longo dos eixos x, y e z, enquanto os giroscópios informam a velocidade angular em torno dos mesmos eixos [19].

É possível utilizar estes sensores para reconhecimento de atividades humanas através de técnicas de aprendizagem de máquina, também conhecida como inteligência artificial, com a qual programas de computador são capazes de reconhecer padrões em dados. No entanto, para que esse reconhecimento seja realizado é necessário que antes o computador "aprenda" com estes dados. Esse aprendizado é realizado na etapa de treinamento, onde o classificador, ou o algoritmo de inteligência artificial utilizado, analisa os dados de modo a reconhecer padrões. É possível que esse reconhecimento seja feito de maneira completamente autônoma por parte do computador, num procedimento conhecido como aprendizagem não supervisionada, ou através de auxílio humano, através da aprendizagem supervisionada [20].

De maneira simplificada, o desafio do algoritmo é atribuir à um vetor X uma classe. Esta classe deve ser escolhida dentre as existentes em um vetor Y. Surge daí, portanto, o termo classificador. Na aprendizagem supervisionada parte da classificação é feita de maneira manual e o algoritmo tenta encontrar, nos dados já classificados, os padrões que irão auxiliar na classificação dos dados restantes, que ainda não foram classificados.

O processo de aprendizagem supervisionada é utilizado tanto para o reconhecimento das atividades através dos dados do acelerômetro, quanto no reconhecimento da localização através da leitura das intensidades de redes Wi-Fi próximas. No caso do reconhecimento das atividades, antes que o algoritmo tente reconhecer qualquer atividade de maneira autônoma, os dados dos sensores dos smartphones são classificados manualmente num procedimento que é apresentado em detalhes na Subseção 3.2.1. Já no caso da localização, o treinamento do algoritmo também é realizado num procedimento que é detalhado na Subseção 3.1.2.

No quesito escolha dos equipamentos, faz-se necessário destacar, no entanto, que a opção de utilização de equipamentos de baixo custo, apesar de diminuir sua barreira para implementação, afeta a precisão final do sistema, uma vez que os modelos de acelerômetro e giroscópio utilizados nos smartphones podem variar consideravelmente de acordo com o modelo do aparelho.

A integração dos dados de localização e do reconhecimento de atividades é a aplicação proposta para fornecer informações referentes ao nível de atividade de trabalhadores ou, mais precisamente, a porcentagem de tempo gasto diretamente na execução de uma determinada tarefa. A determinação do tempo gasto diretamente na execução de uma tarefa, como já mencionado anteriormente, é um dos objetivos da técnica de *work sampling*.

O funcionamento do sistema depende da execução de três etapas:

1. Identificação do local onde a tarefa será realizada:

Nesta etapa, no ambiente onde se dará a execução da tarefa a ser analisada, deverão ser identificadas 3 zonas, sendo uma onde o trabalho será executado de fato, uma para armazenamento de material e outra especificamente designada para descanso;

2. Identificação de Atividades e Amostragem:

Nesta etapa a execução das atividades será analisada através do cruzamento de informações de localização e movimentação. Essas informações são enviadas ao servidor que determina o estado do trabalhador, podendo este ser: "Parado", "Movimentando-se" e "Caminhando";

3. Análise de Produtividade:

Nesta etapa a informação do tempo gasto em uma das zonas especificadas anteriormente é cruzada com estado do trabalhador, determinado na etapa anterior, gerando novos estados, sendo estes "Trabalhando", "Deslocando-se", "Manipulando Material" e "Ocioso". A partir destes dados é apresentado um relatório com a composição de tempo de acordo com cada um destes estados;

3.1.1 Sistema de Posicionamento Interno com base em redes Wi-Fi

O sistema de posicionamento interno adotado para este protótipo é baseado em um *framework*, ou biblioteca, de código aberto que utiliza RSSIs de redes Wi-Fi próximas, aplicando a técnica de *fingerprinting* em conjunto com o Classificador de Naive Bayes para determinação da localização do usuário.

Para compreensão do funcionamento de um sistema de posicionamento interno é necessário também conhecer um pouco mais sobre os sistemas de posicionamento global (GPS). Estes têm se tornado cada vez mais populares no mundo moderno, podendo ser encontrados nos mais diversos tipos de equipamentos, desde de carros a relógios inteligentes. No entanto, esta tecnologia possui limites e não pode ser usada em ambientes internos, onde existam a interferência de paredes, telhados ou qualquer obstáculo capaz de funcionar como barreira para o recebimento dos sinais de satélite utilizados para a determinação da localização dos equipamentos de GPS [21].

Os sistemas de posicionamento interno (IPS), portanto, surgem como uma alternativa capaz de fornecer informações de localização onde os sistemas de GPS não são capazes de chegar, baseando se, mais comumente, na utilização de sinais de rádio, ultrassom e infravermelho [21].

A utilização de ondas de rádio como base para os sistemas de localização interna é, em comparação aos sistemas que utilizam ultrassom e infravermelho, economicamente mais eficiente e de implementação mais simplificada [5].

Dentro da possibilidade de utilização de ondas de rádio como solução para a localização em ambientes internos, duas tecnologias amplamente utilizadas atualmente destacam-se por sua simplicidade de implementação e ampla disponibilidade, sendo elas a de Wi-Fi e Bluetooth.

Ambos os sistemas já tiveram sua aplicabilidade demonstrada tanto nas formas de soluções comerciais quanto acadêmicas. Na construção civil, especificamente, um estudo de caso realizado por Woo et. al, na construção de um túnel em Guangzhou, China, demonstrou a viabilidade da aplicação de um sistema de localização baseado em Wi-Fi e na técnica de *fingerprinting*, apresentada na seção 3.1.2 deste trabalho. Tal sistema, mesmo diante das condições adversas observadas no canteiro de obras, obteve precisão de localização com margem de erro de 5m.

3.1.2 Fingerprinting

A tecnologia de *fingerprinting* é a mais utilizada por sistemas de localização interna tanto para sistemas baseados em Wi-Fi, quanto para sistemas baseados em Bluetooth. Tal tecnologia consiste na obtenção de uma coleção de "impressões digitais", ou *fingerprints*, compostas por um conjunto de intensidade de sinais de rádio mensurados em um determinado local. Este conjunto de sinais, mensurados através das RSSIs, tende a compor uma combinação única, que varia de local para local [15].

Os sistemas que utilizam a tecnologia de *fingerprinting* dividem seu funcionamento em duas etapas, a de treinamento e a de localização propriamente dita. Na fase de treinamento, os seguintes passos são executados:

1. São geradas combinações únicas de RSSIs que dão origem as “impressões digitais” de cada local;
2. As combinações são associadas a uma identificação dada a localização atual;
3. Os dados contendo a combinação de RSSIs e a localização são armazenados em um banco de dados;

Na fase de localização, o sistema segue os seguintes passos:

1. As impressões digitais são novamente obtidas com base na combinação de RSSIs;
2. As impressões digitais são comparadas às armazenadas no banco de dados;
3. Um modelo estatístico classifica as probabilidades das impressões digitais associadas às localizações armazenadas corresponderem a localização atual baseando-se na utilização de um algoritmo conhecido como Naive Bayes;
4. O sistema, então, escolhe o local que apresenta a maior probabilidade como correspondente ao atual e assim determina a localização do usuário;

A arquitetura do sistema de localização interna é demonstrada na Figura 1.

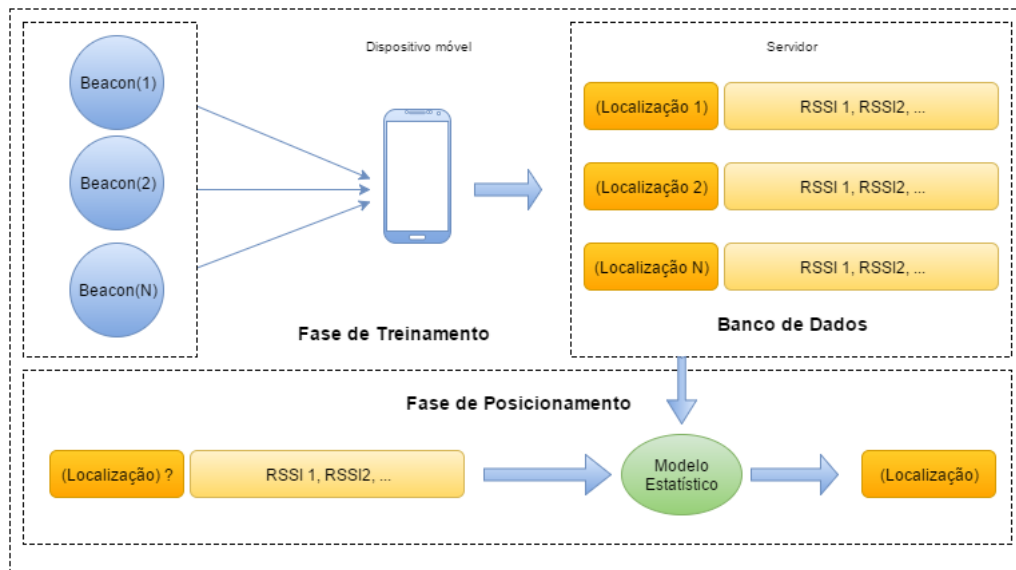


Figura 1: Ilustração de modelo de IPS utilizando a técnica de *fingerprinting*.

3.1.2.1 Classificador de Naive Bayes

O classificador de Naive Bayes foi desenvolvido à partir da implementação do Teorema de Bayes que com base em uma variável x retorna a probabilidade de C , como apresentado na Equação 3.1

$$p(C_k|x) = \frac{p(C_k)p(x|C_k)}{p(x)} \quad (3.1)$$

Aplicando-se à técnica de *fingerprinting*, o Teorema de Bayes dá a probabilidade de que um *fingerprint* x pertença à classe C , de modo que a classe C representa a identificação de uma determinada localização e x é um vetor que contém as leituras dos RSSIs das redes Wi-Fi próximas [3].

Aplica-se o termo "ingênuo" (*naive*) ao classificador devido ao fato de que este classificador assume que os valores dos vetores de entrada são todos independentes entre si. Ao fazer essa suposição o cálculo da probabilidade condicional $P(x|C)$ é facilitado, o que simplesmente se transforma no produto da probabilidade de cada elemento do vetor x dada uma classe C :

$$\prod_i P(x_i|C) \quad (3.2)$$

Considera-se que a *evidência*, $P(x)$, serve apenas para normalização dos resultados e, portanto, é geralmente ignorada e a equação é reduzida para:

$$P(C|x) = P(C) \prod_i P(x_i|C) \quad (3.3)$$

Para classificar x , é utilizada a regra de decisão de Máximo a Posteriori (*Maximum a Posteriori*, MAP), que define que a fórmula anterior seja computada para cada classe de modo que seja selecionada aquela que possua a probabilidade mais alta como resultado:

$$\text{escolha } C_i \text{ se } P(C_i|x) = \max_k P(C_k|x) \quad (3.4)$$

Esta etapa assegura que o classificador escolherá o resultado correto contanto que a classe correta possua a probabilidade mais alta, mesmo que as probabilidades não estejam distribuídas corretamente entre as classes [3].

3.2 SISTEMA DE DETECÇÃO DE MOVIMENTOS COM BASE EM DADOS DE ACELERÔMETRO E GIROSCÓPIO

O sistema de detecção de movimentos aplicado a este protótipo resume-se na identificação automática de três atividades primárias: "Parado", "Caminhando" e "Movimentando-se". Para tal, os dados de acelerômetro e giroscópio de um *smartphone* são coletados e classificados manualmente, para que posteriormente sejam enviadas a um programa aqui identificado como classificador.

3.2.1 Coleta de Dados

Os dados de acelerômetro e giroscópio são coletados de um *smartphone*, com sistema operacional Android. Estes dados são capturados em um taxa aproximada de 50Hz e enviados a um servidor desenvolvido na linguagem Python operando sobre o sistema operacional Linux.

O *smartphone* é fixado ao cinto do usuário na posição horizontal, como demonstra a Figura 2.



Figura 2: Demonstração do posicionamento do *smartphone* no corpo do usuário.

A escolha do local para fixação do *smartphone* é baseada no fato de que a região da cintura está localizada próximo ao centro de gravidade do corpo captando, assim, grande parte

de seu movimento sem que haja, no entanto, interferência de movimentos involuntários ou de pouca relevância para esta aplicação, como ocorre com o posicionamento na região do braço e pernas [22].

Além da escolha do posicionamento na região da cintura, a fixação ao cinto reduz a quantidade de interferência nos dados provocadas por movimentação irregular do *smartphone* ocorrida, por exemplo, caso este estivesse solto no bolso do usuário. É possível, contudo, utilizar o *smartphone* desta forma para captura de dados, sendo esta considerada até mesmo menos intrusiva. Porém, tal forma aumenta a complexidade da limpeza e análise dos dados, requerendo a possível aplicação de mais filtros e extração de mais características para composição de vetores. Surge, portanto, a possibilidade de estudos futuros com foco na comparação da precisão dos resultados de testes com o *smartphone* posicionado de diferente modos no usuário.

3.2.2 Classificação Manual

Um usuário utilizando um *smartphone* acoplado ao cinto, como apresentado anteriormente, executa um serviço de alvenaria enquanto é gravado em vídeo. Posteriormente, utilizando um editor de planilhas e um editor de vídeo, os dados obtidos do acelerômetro são sincronizados ao vídeo e classificados manualmente, indicando na planilha qual atividade está sendo executada no instante em que cada amostra é capturada pelos sensores. O processo é ilustrado pela Figura 3.

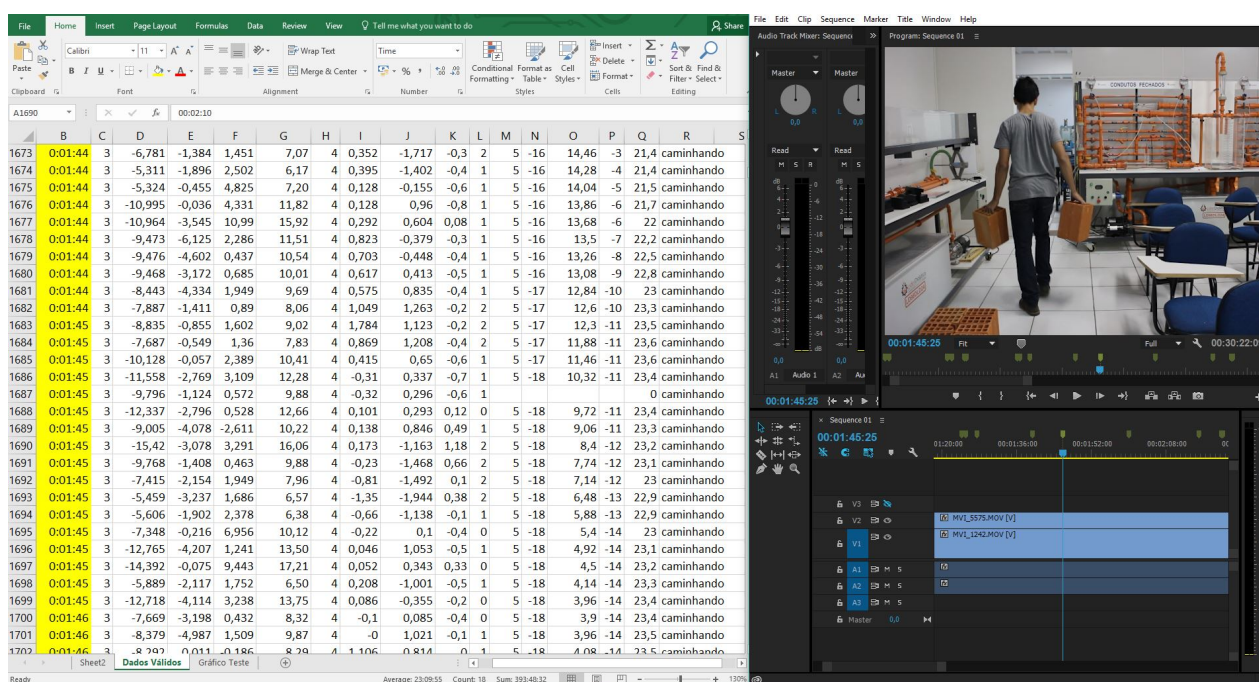


Figura 3: Demonstração do processo de classificação dos dados dos sensores.

As atividades são definidas como "Parado", "Caminhando" e "Movimentando-se". Entende-

se aqui "Parado", como estar parado, de pé, sentado ou agachado, sendo estas situações executadas na fase de treinamento; "Movimentando-se" é definido como estar executando movimentos relativos a execução de um trabalho qualquer ou de manuseio de material sem deslocar-se de um lugar para outro, excluindo-se, portanto, a classe "Caminhando", que é atribuída, justamente, quando um usuário se desloca de um lugar a outro.

Após a classificação manual, os dados são exportados no formato CSV (*Comma Separated Values*), permitindo que sejam processados, para eliminação de ruídos dos sensores, e avaliados pelo classificador no servidor. É importante ressaltar que o processo de classificação manual só é realizado uma única vez.

3.2.3 Análise e limpeza de dados

Os dados coletados diretamente dos sensores do *smartphone* estão sujeitos a interferência de diferentes naturezas. Cabe, portanto, a aplicação de filtros de modo que essas interferências sejam mitigadas ou, em alguns casos, completamente eliminadas.

Para esta aplicação adotou-se o modelo proposto por Anguita et al, onde aplica-se, inicialmente, um filtro linear de suavização, mais especificamente um filtro de mediana. Este filtro tem sua implementação dada pela simples obtenção da mediana das matrizes unidimensionais, ou vetores, obtidas à partir de uma janela amostral de tamanho 3 aplicada sobre a série temporal composta pelos dados referentes a aceleração linear e angular dos eixos X, Y e Z, colhidos do acelerômetro e giroscópio, respectivamente.

Posteriormente, aplica-se a estes mesmos dados um filtro do tipo passa-baixa, cuja função é, por definição, eliminar o sinal das frequências superiores à faixa de corte definida e manter as inferiores. É proposta, mais especificamente, a utilização de um filtro de Butterworth de terceira ordem, cuja implementação se dá a partir da Equação 3.5.

$$|H(\omega)|^2 = \frac{1}{1 + (\omega/\omega_c)^{2N}} \quad (3.5)$$

Onde:

n é a ordem do filtro;

ω_c é a frequência de corte;

e ω é a frequência angular do sinal em radianos por segundo.

Anguita et. al, propõe uma frequência de corte na faixa de 20 Hz baseando-se no fato de que esta faixa é suficiente para capturar o movimento humano e afirmando que 99% de sua energia está contida em uma faixa inferior a 15 Hz.

Em seguida, o sinal do acelerômetro, que contém componentes de movimento corporal e gravitacional, é separado usando, novamente, um filtro passa-baixa de Butterworth. A frequência

de corte desta vez é definida na faixa de 0.3 Hz, sendo as frequências inferiores à estas, referentes à aceleração gravitacional [22].

Após a aplicação dos filtros, sinais adicionais são obtidos à partir da determinação dos módulos dos vetores e derivadas temporais (arranque da/dt e aceleração angular dw/dt).

Finalmente, são obtidas amostras dos sinais no domínio do tempo, com janelas amostrais de tamanho fixo com comprimento equivalente a 2.56s e sobreposição de 50% entre eles. Estas escolhas, são suficientes para que a cadência média de uma pessoa caminhando seja reconhecida corretamente pelo classificador [22]. Os sinais também são mapeados no domínio da frequência aplicando-se o algoritmo de Transformada Rápida de Fourier, cuja implementação baseia-se na Equação 3.6.

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi k n/N} \quad (3.6)$$

assumindo-se X_0, \dots, X_{N-1} como números complexos e $k = 0, \dots, N - 1$

Portanto, à partir deste método, 17 sinais são obtidos, como mostrado na Tabela 1. Estes sinais irão compor os vetores de características, também conhecidos como vetores espectrais, passados ao classificador nas fases de treinamento e classificação.

Tabela 1: Sinais no domínio de tempo e frequência obtidos à partir dos sensores de smartphone

Nome	Tempo	Frequência
Aceleração Corporal	X	X
Aceleração Gravitacional	X	
Arranque da Aceleração Corporal	X	X
Velocidade Angular Corporal	X	X
Aceleração Angular Corporal	X	
Módulo da Aceleração Corporal	X	X
Módulo da Aceleração Gravitacional	X	
Módulo do Arranque da Aceleração Corporal	X	X
Módulo da Velocidade Angular Corporal	X	X
Módulo da Aceleração Angular Corporal	X	X

3.2.4 Classificador de Árvore de Decisão

O classificador adotado para este protótipo é conhecido como Classificador de Árvore de Decisão (*Decision Tree Classifier*, DTC). Sua escolha foi dada pela simplicidade de implementação e resultados satisfatórios na aplicação de reconhecimento e classificação de movimentos, como demonstrados por vários pesquisadores, dentre eles [23, 24].

O método de classificação baseado em árvore de decisão utilizado neste trabalho é o de Árvore de Regressão e Classificação (*Classification and Regression Tree*, CART). O método

CART divide os exemplos de treinamento no espaço espectral em regiões retangulares, também conhecidas como "nós", e atribui uma classe a cada região. O processo inicia-se com todas as classes distribuídas ao longo do espaço espectral e examina todas as possibilidades de divisões binárias em cada característica que compõe o espaço espectral [25].

Subsequentemente, uma subdivisão é selecionada, caso possua o melhor critério de otimização, dando origem ao índice de diversidade de Gini, como demonstrado pela Equação 3.7.

$$I_g(f) = 1 - \sum_{i=1}^k f_i^2 \quad (3.7)$$

Onde:

I_g é o Índice de Gini;

f_i é a fração de itens marcados com o valor i ;

e k é o número de classes.

Enquanto todos os "nós" não forem considerados "puros" o processo irá se repetir. Para considerar que um "nó" é "puro", ele deve conter apenas observações de uma única classe, de modo que passa-se a assumir que o Índice de Gini é zero, ou que há um número inferior a 10 observações a serem divididas [23].

3.2.5 Treinamento do Classificador

O treinamento do classificador é realizado através do fornecimento de vetores compostos pelos sinais obtidos após o processamento apresentado nas seções anteriores, e das classes das quais estes vetores pertencem. No caso deste trabalho, as classes utilizadas foram as referentes às atividades mencionadas na Seção 3.2.1 ("Parado", "Caminhando" e "Movimentando-se").

Para verificação da eficiência do classificador é necessário que parte dos dados sejam destinadas ao treinamento do classificador e parte para teste. Deste modo, 70% dos dados foram destinados ao treinamento e 30% para o teste de precisão do algoritmo, cujos resultados são apresentados no próximo capítulo.

3.3 INTERSEÇÃO DOS DADOS DE MOVIMENTAÇÃO E LOCALIZAÇÃO

Para gerar as classes finais ("Trabalhando", "Deslocando-se", "Manipulando Material" e "Ocioso"), o sistema protótipo cruza as classes iniciais ("Parado", "Caminhando" e "Movimentando-se") com as informações de localização seguindo um fluxo de condições conforme demonstrado pela Tabela 2.

Tabela 2: Avaliação das classificações iniciais para determinação das classificações finais

Localização do Usuário	Classificação Primária	Classificação Final
Zona de Armazenamento	Parado	Ocioso
	Movimentando-se	Manipulando Material
	Caminhando	Deslocando-se
Zona de Descanso	Parado	Ocioso
	Movimentando-se	Ocioso
	Caminhando	Deslocando-se
Zona de Execução	Parado	Ocioso
	Movimentando-se	Trabalhando
	Caminhando	Deslocando-se

3.4 APRESENTAÇÃO DAS BIBLIOTECAS UTILIZADAS

As bibliotecas utilizadas para processamento de dados neste trabalho fornecem uma implementação simplificada de grande parte dos algoritmos aqui mencionados e estão disponíveis em código aberto para a linguagem de programação Python.

3.4.1 Scikit-learn

A *scikit-learn* é uma biblioteca de código aberto que reúne um conjunto de algoritmos de aprendizagem de máquina. Dentre eles o mais importante para a análise dos dados utilizados neste trabalho é a implementação do classificador de árvore de decisão. O código utilizado no programa desenvolvido para este trabalho encontra-se no Apêndice A.

3.4.2 Pandas

Pandas é uma biblioteca de código aberto, também desenvolvida para a linguagem Python, amplamente utilizada em análise de dados. Ela fornece um conjunto de ferramentas, destacando como as mais importantes para este trabalho as implementações de ferramentas para manipulação de dados no formato CSV (*Comma Separated Values*) e o tratamento de séries de dados incompletos.

3.4.3 NumPy

Numpy é uma biblioteca de código aberto para a linguagem Python que provê suporte para operações matemáticas de diferentes complexidades, como processamento de matrizes multidimensionais. Esta função se destaca como a mais importante para este trabalho, uma vez que os dados em formato de tabela utilizados neste trabalho são transformados em matrizes para que possam ser processados pelos filtros apresentados na seção de análise e limpeza de dados.

3.4.4 SciPy

Assim como as anteriores, *SciPy* é uma biblioteca de código aberto desenvolvida na linguagem Python, que fornece acesso a implementação de diversos algoritmos utilizados em computação científica. Dentre eles destacam-se, especialmente, os algoritmos de processamento de sinais, utilizados na implementação dos filtros passa-baixa aplicados na limpeza dos dados capturados diretamente do *smartphone* e na aplicação da Transformada Rápida de Fourier para mapeamento de parte dos dados no domínio das frequências, como apresentado na Tabela 1.

3.4.5 Matplotlib

A biblioteca *Matplotlib* é considerada uma extensão da biblioteca *NumPy* e fornece ferramentas para criação de gráficos de diferentes tipos. Neste trabalho ela é utilizada para a criação da "Matriz de Confusão" apresentada na seção de resultados.

3.5 APRESENTAÇÃO DO EXPERIMENTO PROPOSTO

Para treinar e posteriormente testar o protótipo, é proposta a simulação da execução de um serviço de alvenaria dentro de um ambiente controlado. A execução consiste na construção de uma parede de $1m^2$, com o objetivo de capturar os padrões de movimento relativos a execução deste serviço de modo natural. O ambiente em questão refere-se às proximidades da instalação do laboratório de materiais de construção das Faculdades Integradas de Caratinga. O ambiente é dividido em três zonas, sendo uma dedicada ao armazenamento de material, uma para descanso e uma para execução do serviço propriamente dito, como demonstrado nas Figuras 4, 5 e 6, respectivamente. A zona definida como zona de armazenamento é a única situada em um ambiente fechado, separada por uma parede dos outros ambientes e localizada dentro das instalações do laboratório de hidráulica. A zona de execução encontra-se em local parcialmente aberto, dentro do laboratório de materiais de construção. A zona de descanso foi definida em local completamente aberto, próximo ao laboratório de materiais de construção.



Figura 4: Local designado para almacenamiento de material.



Figura 5: Local designado para descanso.



Figura 6: Local designado para execução do trabalho.

Para captura dos dados um *smartphone* é fixado à cintura de um usuário, como já demonstrado na Figura 2, e o serviço de alvenaria é executado enquanto os dados referentes à localização, ao acelerômetro e ao giroscópio são enviados ao servidor. Todo processo é capturado em vídeo e executado pelo período de 30 minutos.



Figura 7: Execução do Serviço de Alvenaria.

4 RESULTADOS E DISCUSSÕES

4.1 PRECISÃO DO CLASSIFICADOR

Para avaliar a precisão do classificador os dados capturados são analisados utilizando a biblioteca *scikit-learn*. A análise da precisão é realizada separando-se parte dos dados já pré-processados para comporem um conjunto denominado "conjunto de teste". Este conjunto é passado ao classificador sem as classificações que foram realizadas manualmente em uma etapa anterior, como demonstrado na Seção 3.2.2. Cabe, portanto, ao classificador fornecer essas classes de maneira automatizada para que possam ser comparadas com as classificações manuais. Ou seja, comparam-se as classes fornecidas pelo classificador com as classes esperadas.

A precisão aqui é representada, portanto, pela taxa acertos do classificador. Matematicamente essa taxa pode ser expressa pela Equação 4.1.

$$P = \frac{N_{ic}}{N_{it}} * 100 \quad (4.1)$$

Onde:

N_{ic} é o número de instâncias de classe previstas corretamente pelo classificador;
e N_{it} é o número total de instâncias naquela classe.

A partir desta análise as precisões obtidas para cada uma das classes analisadas foram:

- Parado: 85%
- Movimentando: 75%
- Caminhando: 96%
- Média: 85%

Além dos dados relativo a precisão em cada uma das classes, foi gerada uma "matriz de confusão", que apresenta em um dos eixos a atividade correta, sendo a esperada como resultado do classificador, e em outro a atividade prevista pelo classificador.

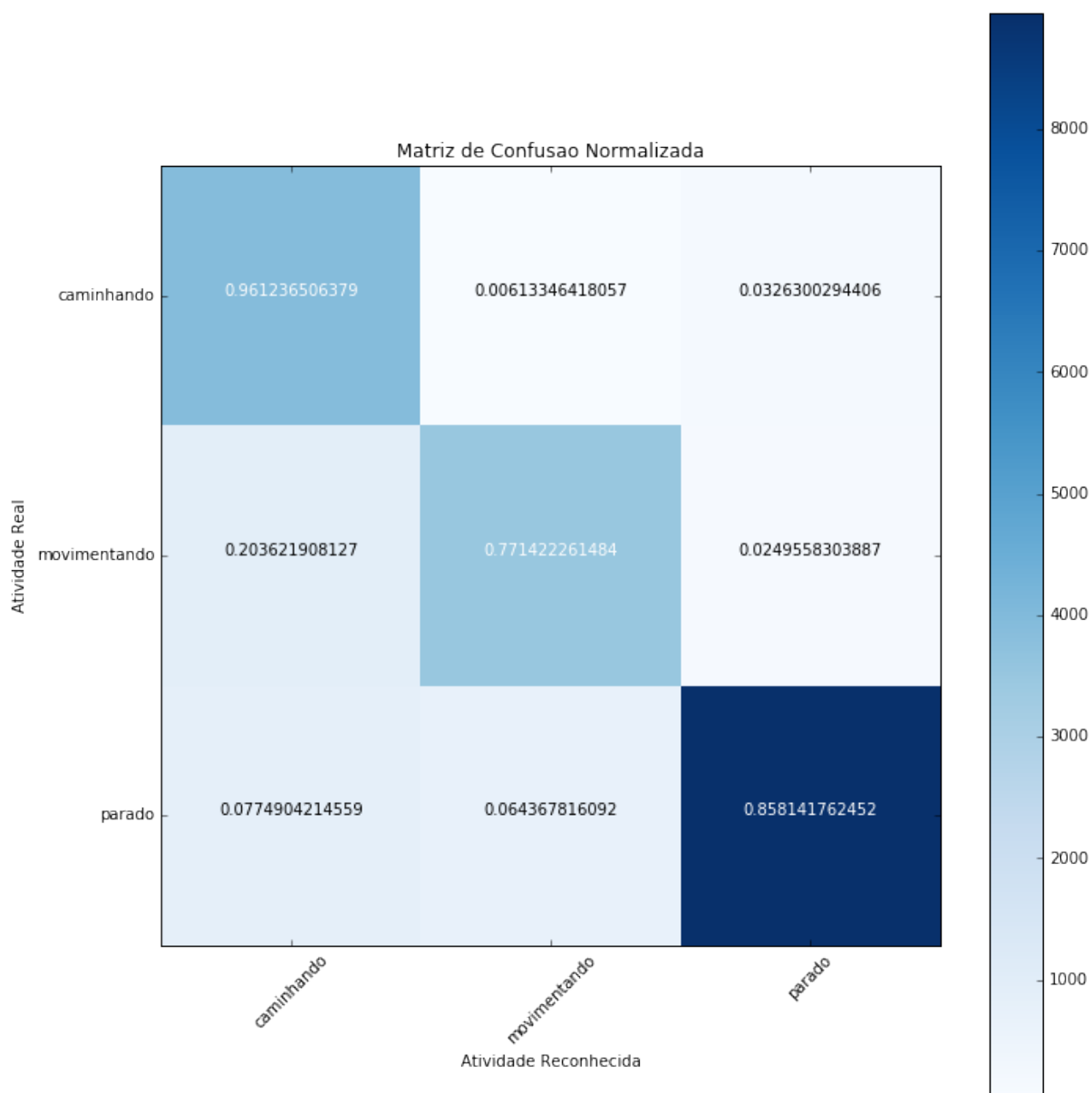


Figura 8: Matriz de Confusão Normalizada.

A partir da matriz de confusão normalizada, demonstrada na Figura 8, observa-se, por exemplo, que a classe "movimentando" foi, em 20% das vezes em que era esperada, confundida com a classe "caminhando". O inverso só aconteceu, no entanto, em 0,6% das vezes. Isso ocorre devido ao fato de que, uma vez que o padrão de caminhar é reconhecido, ele se torna imediatamente evidente ao classificador. Já a classe "movimentando", por não ser específica, englobando uma série de possíveis movimentos, pode ser inicialmente confundida com a classe "caminhando". Contudo, para o propósito deste trabalho essa confusão é aceitável.

4.2 PRECISÃO DO SISTEMA DE LOCALIZAÇÃO INTERNA

A precisão do Sistema de Localização Interna foi avaliada comparando-se as localizações reais, registradas em vídeo, com as localizações informadas pelo sistema, obtendo diferentes níveis de precisão para cada localização. A análise foi realizada de maneira semelhante a realizada para o classificador apresentado anteriormente. A precisão abaixo representa, portanto, a taxa de acertos do sistema de localização interna para cada uma das zonas delimitadas para este trabalho.

- Zona de Execução: 92%
- Zona de Armazenamento: 95%
- Zona de Descanso: 89%
- Média: 92%

A variação na precisão da identificação correta de cada local, ocorre devido a diversos fatores, dentre eles:

- Geometria do Local;
- Disponibilidade de Sinais de Redes Wi-Fi;
- Atraso na identificação do local.

4.3 PRECISÃO DAS CLASSIFICAÇÕES FINAIS

Como já mencionado anteriormente, o objetivo final do sistema é classificar a utilização do tempo do trabalhador através da análise de nível de atividade, indicando o tempo dedicado a manipulação de material, ao deslocamento e à execução do trabalho em si.

Para chegar a composição da classificação final utiliza-se a Tabela 3. Esta tabela apresenta o fluxo de condições aplicado aos dados classificados manualmente e aos dados com as classes previstas pelo classificador. A implementação foi realizada na linguagem Python, com auxílio da biblioteca *Pandas*. A Imagem 9 apresenta o fluxograma do sistema utilizado para obtenção das classificações finais.

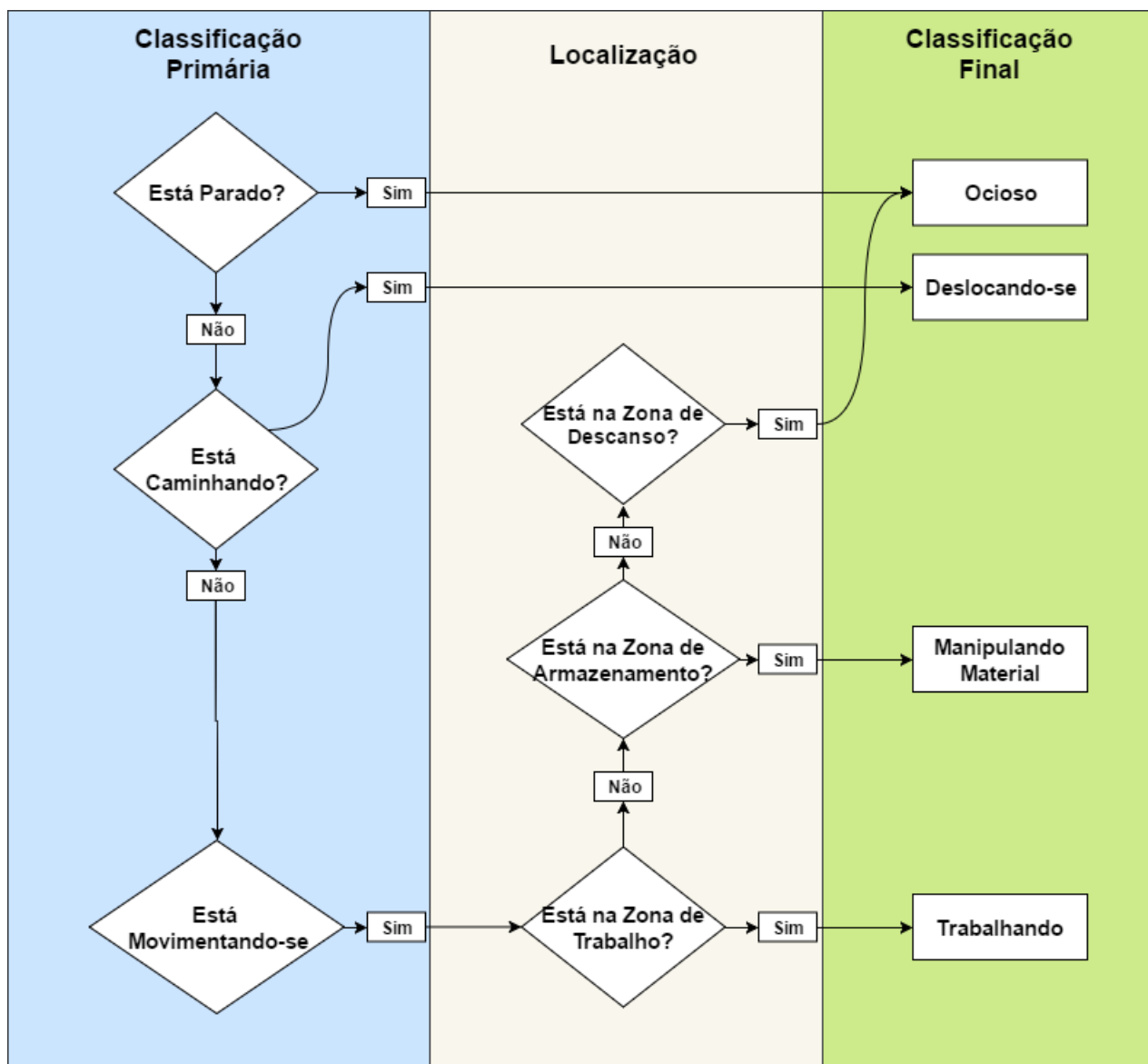


Figura 9: Fluxograma do Sistema de Classificação Final.

Os resultados são apresentados na Tabela 3. Os dados gerados pelo sistema foram salvos no formato CSV e importados em um editor de planilhas. Cada observação registrada representa o equivalente a 0,02s (considerando a taxa de amostragem de 50 Hz utilizada na captura dos dados). Portanto, o tempo dedicado a cada uma das classes finais foi obtido a partir destes cálculos.

Tomando a classificação manual como parâmetro para estimar a precisão do sistema, observa-se que a variação máxima ocorreu na classificação do deslocamento e a mínima no estado ocioso. Contudo há grande diferença nas amostras para cada uma das classificações, sendo necessário, portanto, a avaliação do desempenho com distribuições de tempo mais equilibradas. Deste modo, a opção por capturar as amostras durante a execução de um serviço de alvenaria de modo natural, sem que houvesse tentativas deliberadas de distribuir o tempo de maneira mais

Tabela 3: Comparação das Classificações Finais Manuais com Base em Vídeo e a Classificação Automática Realizada Pelo Sistema.

Classificação Final	Nº de Observações	Composição do Tempo de Acordo Com a Avaliação Manual		Composição do Tempo de Acordo com o Sistema Protótipo		Variação Entre a Classificação Manual e Classificação do sistema	
		%	Tempo (min)	%	Tempo (min)	%	Tempo (min)
Trabalhando	61620	70%	20,54	63%	18,49	7%	2,05
Deslocando-se	11444	13%	3,81	28%	7,04	15%	-3,23
Ocioso	8803	10%	2,93	8%	2,35	2%	0,59
Manipulando Material	6162	7%	2,05	1%	1,47	6%	0,59
TOTAL	88029	100%	29,34	100%	29,34	Média	1,61

equilibrada mostrou-se pouco eficiente para estimativa da precisão final do classificador.

Ao avaliar a precisão do sistema partindo da análise separada de cada um de seus componentes (Sistema de Localização Interna e Sistema de Reconhecimento de Movimentos) os resultados mostram-se promissores. Cabe, portanto, a implementação de estudos futuros dedicados a análises isoladas das aplicações de cada um destes sistemas na construção civil e na utilização de outros métodos para avaliação do desempenho da integração destes dois sistemas e de sua implementação no setor construção civil.

5 CONCLUSÃO

Este trabalho avaliou a possibilidade da implementação de tecnologias modernas e de baixo custo para a análise automatizada de atividades realizadas na construção civil. Este processo, capaz de fornecer informações importantes sobre o uso do tempo dos trabalhadores, foi avaliado através de experimentos conduzidos em ambiente controlado, onde as atividades relacionadas a execução de um serviço de alvenaria foram realizadas. Os resultados obtidos demonstraram que a tecnologia automatizada para reconhecimento de atividades, a tecnologia de localização interna e a união destas duas para análise de atividades demonstrou-se promissora. No entanto, cabe ainda a possibilidade de aprimoramento do sistema utilizado a fim de que haja mais precisão nos resultados antes de que seja considerada sua implementação em um ambiente sob condições reais de trabalho.

Uma das limitações atuais mais notórias é relativa ao equipamento utilizado, neste caso um smartfone. A variação na precisão dos resultados fornecidos pelos acelerômetros e giroscópios de cada modelo de smartfone pode influenciar consideravelmente os resultados finais no caso de uma implementação em maior escala. Além disso, a forma como o smartfone foi utilizado pode ser considerada intrusiva, afetando, mais uma vez, a precisão dos resultados. O método utilizado para avaliação da precisão do sistema também mostrou-se pouco eficiente ao gerar um número reduzido de observações para avaliação de algumas das classificações.

Contudo, as técnicas e tecnologias utilizadas para processamento dos dados fornecidos pelo smartfone demonstraram-se eficientes. Fica aberta, portanto, a possibilidade de estudos futuros com a utilização destas mesmas tecnologias em conjunto com equipamentos desenvolvidos especificamente para este propósito, resultando em um dispositivo com dimensões menores tornando-o, conseqüentemente, menos intrusivo. Já há suporte para o desenvolvimento de tal equipamento, uma vez que existem plataformas de baixo custo no mercado, como a plataforma Arduino e a Raspberry Pi. Ambas podem ser equipadas com módulos de Wi-Fi, acelerômetro e giroscópio, sendo estes, como demonstrado neste trabalho, os únicos equipamentos necessários para implementação do sistema de localização interna e do sistema de detecção automatizada de movimentos. Estes dois últimos, portanto, integrados para o fornecimento de uma análise automatizada das atividades realizadas na construção civil.

Além das possibilidades de estudos futuros já mencionadas, também há a possibilidade de estudos voltados apenas ao aprimoramento do sistema de detecção automatizada de movimentos através da realização de testes com outros algoritmos de inteligência artificial, uma coleta mais ampla de dados e com a utilização de mais participantes nos estudos. As aplicações deste sistema na construção civil estendem-se a classificação e ao reconhecimento dos movimentos relacionados a execução dos mais diversos serviços realizados neste setor. Tal sistema poderia auxiliar a automatizar o processo de supervisão de mão de obra e fornecer dados importantes no

gerenciamento de projetos na indústria da construção civil.

6 REFERÊNCIAS BIBLIOGRÁFICAS

- [1] SOUZA, Ubiraci Espinelli Lemes de. Como medir a produtividade da mão-de-obra na construção civil. Encontro Nacional de Tecnologia do Ambiente Construído, v. 8, n. 1, 2000.
- [2] MENDES, Luís Fernando Melo et al. A produtividade da Construção Civil brasileira. Disponível em: <http://www.cbic.org.br/sites/default/files/FINAL1_CBIC_64pgs.pdf>. Acesso em: 26 de maio de 2017.
- [3] BEKKELIEN, Anja; DERIAZ, Michel; MARCHAND-MAILLET, Stéphane. Bluetooth indoor positioning. Master's thesis, University of Geneva, 2012.
- [4] CURRAN, Kevin et al. An evaluation of indoor location determination technologies. Journal of Location Based Services, v. 5, n. 2, p. 61-78, 2011.
- [5] WOO, Sunkyu et al. Application of WiFi-based indoor positioning system for labor tracking at construction sites: A case study in Guangzhou MTR. Automation in Construction, v. 20, n. 1, p. 3-13, 2011.
- [6] AGGARWAL, Jake K.; RYOO, Michael S. Human activity analysis: A review. ACM Computing Surveys (CSUR), v. 43, n. 3, p. 16, 2011.
- [7] SOUZA, Ubiraci Espinelli Lemos de. Como aumentar a Eficiência da Mão-de-Obra. São Paulo, SP. PINI, 2006.
- [8] MESSA, Alexandre. Indicadores de produtividade: uma breve revisão dos principais métodos de cálculo. IPEA. Radar: tecnologia, produção e comércio exterior, 28, pp. 17-26, 2013.
- [9] MACEDO, Mariano de Matos. Gestão da produtividade nas empresas. Revista Organização Sistêmica, v. 1, n. 1, p. 110-119, 2012.
- [10] NETO, Jorge Batlouni. Prefácio. In: SOUZA, Ubiraci Espinelli Lemos de. Como aumentar a Eficiência da Mão-de-Obra. São Paulo, SP. PINI, 2006.
- [11] SOUZA, Ubiraci Espinelli Lemos de. Produtividade e custos dos sistemas de vedação vertical. Tecnologia e gestão na produção de edifícios: vedações verticais. 1998. In: SOUZA, Ubiraci Espinelli Lemes de. Como medir a produtividade da mão-de-obra na construção civil. Encontro Nacional de Tecnologia do Ambiente Construído, v. 8, n. 1, 2000.
- [12] CHENG, Tao et al. Automated task-level activity analysis through fusion of real time location sensors and worker's thoracic posture data. Automation in Construction, v. 29, p. 24-39, 2013.
- [13] CRAWFORD, Paul; VOGL, Bernard. Measuring productivity in the construction industry. Building Research and Information, v. 34, n. 3, p. 208-219, 2006.
- [14] ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. Measu-

- ring Productivity: Measurement of Aggregate and Industry-level Productivity Growth: OECD Manual. Organization for Economic Co-operation and Development, 2010.
- [15] PARK, Hee-Sung; THOMAS, Stephen R.; TUCKER, Richard L. Benchmarking of construction productivity. *Journal of Construction Engineering and Management*, v. 131, n. 7, p. 772-778, 2005.
- [16] LIOU, Fwu-Shiun; BORCHERDING, John D. Work sampling can predict unit rate productivity. *Journal of Construction Engineering and Management*, v. 112, n. 1, p. 90-103, 1986.
- [17] PEINADO, Jurandir; GRAEML, Alexandre Reis. *Administração da produção. Operações industriais e de serviços*. Unicenp, 2007.
- [18] THOMAS, H. Randolph; DAILY, Jeffrey. Crew performance measurement via activity sampling. *Journal of Construction Engineering and Management*, v. 109, n. 3, p. 309-320, 1983.
- [19] JOSHUA, Liju; VARGHESE, Koshy. Accelerometer-based activity recognition in construction. *Journal of computing in civil engineering*, v. 25, n. 5, p. 370-379, 2010.
- [20] KOTSIANTIS, Sotiris B.; ZAHARAKIS, I.; PINTELAS, P. *Supervised machine learning: A review of classification techniques*. 2007.
- [21] GREENEMEIER, Larry. A positioning system that goes where GPS can't. *Scientific American*, Jan, v. 23, 2008.
- [22] ANGUIA, Davide et al. A Public Domain Dataset for Human Activity Recognition using Smartphones. In: ESANN. 2013.
- [23] AKHAVIAN, Reza; BEHZADAN, Amir H. Smartphone-based construction workers' activity recognition and classification. *Automation in Construction*, v. 71, p. 198-209, 2016.
- [24] PHAM, Cuong; OLIVIER, Patrick. SliceAndDice: Recognizing food preparation activities using embedded accelerometers. In: *European Conference on Ambient Intelligence*. Springer Berlin Heidelberg, 2009. p. 34-43.
- [25] BISHOP, Christopher M. *Pattern recognition. Machine Learning*, v. 128, p. 1-58, 2006.

APÊNDICE A PROGRAMA DE PRÓPRIA AUTORIA PARA PROCESSAMENTO DOS SINAIS RECEBIDOS PELO *SMARTPHONE* E RECONHECIMENTO DE ATIVIDADES

Programa de própria autoria para processamento dos sinais recebidos pelo *smartphone* e reconhecimento de atividades.

```
# In[52]:

from scipy.signal import medfilt, butter, lfilter, freqz
import matplotlib
import cufflinks as cf
import numpy as np
from sklearn import tree
from sklearn.metrics import classification_report
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn import svm
from sklearn.neural_network import MLPClassifier
from sklearn import metrics
import pandas as pd

print "Finalizado"

# In[53]:

# Filtros Butterworth
def butter_lowpass(cutoff, fs, order=3):
    nyq = 0.5 * fs
    normal_cutoff = cutoff / nyq
    b, a = butter(order, normal_cutoff, btype='low', analog=False)
    return b, a

def butter_highpass(cutoff, fs, order=3):
    nyq = 0.5 * fs
    normal_cutoff = cutoff / nyq
    b, a = butter(order, normal_cutoff, btype='high', analog=False)
    return b, a

def butter_lowpass_filter(data, cutoff, fs, order=3):
    b, a = butter_lowpass(cutoff, fs, order=order)
    y = lfilter(b, a, data)
    return y

def butter_highpass_filter(data, cutoff, fs, order=3):
    b, a = butter_highpass(cutoff, fs, order=order)
    y = lfilter(b, a, data)
    return y
```

```
# In[103]:

#Leitura dos dados brutos
raw_data = pd.read_csv("Classificacao-Teste2.csv", header = None, skiprows=1, delimiter = ";",

#Remocao dos dados nulos
data = raw_data.dropna(axis=0,how="any")

#Separacao em fetures e labels (x,y)
x = data[[1,2,3,5,6,7]]
y = data[[12]]

# In[57]:

#calculo do modulo do vetor
def norm(df,x=0,y=1,z=2):
    result = df[[x,y,z]].apply(lambda x: np.sqrt(x.dot(x)),axis = 1)
    return result

# Aplicacao do filtro de mediana nos dados
mfiltered = medfilt(x)

# Aplicacao do filtro Butterwoth Passa-baixa a 20Hz
blfiltered = butter_lowpass_filter(mfiltered, 20, 50, 3)
blfiltered = pd.DataFrame(blfiltered)

#Separacao em Aceleracao(Acc) e Aceleracao Angular (gyr)
acc = blfiltered[[0,1,2]]
gyr = blfiltered[[3,4,5]]

# Aceleracao Gravitacional
grav_acc = butter_lowpass_filter(acc, 0.3, 50, 3)
grav_acc = pd.DataFrame(grav_acc)

#Modulo Aceleracao gravitacion
grav_acc_mag = norm(grav_acc)

#Aceleracao Corporal
body_acc = acc - grav_acc

#Modulo da Aceleracao Corporal
body_acc_mag = norm(body_acc)

#Aceleracao Angular Corporal
body_gyr = gyr

#Modulo de Aceleracao Angular Corporal
body_gyr_mag = norm(body_gyr,3,4,5)
body_gyr_mag

#Uniao dos dados em um unico dataframe
df_list = [body_acc,body_acc_mag,body_gyr,body_gyr_mag,grav_acc,grav_acc_mag]
df = pd.concat(df_list,axis=1)
```



```
df.columns = [i for i in range(df.shape[1])]

# In[101]:

def norm_shape(shape):
    '''
    Normalizacao de Matrizes.
    '''
    try:
        i = int(shape)
        return (i,)
    except TypeError:
        # shape nao e um numero
        pass

    try:
        t = tuple(shape)
        return t
    except TypeError:
        # shape nao e iteravel
        pass

    raise TypeError('shape deve ser int, ou tuple de ints')

from itertools import product

def sliding_window_nd(a,ws,ss = None):
    '''
    Retorna uma janela amostral sobre uma matriz de qualquer dimensao

    Parameters:
        a - uma matriz n dimensional
        ws - tamanho da janela amostral
        ss - um int ou tuple representando o quanto mover a janela
            em cada dimensao
    '''
    if None is ss:
        ss = ws

    ws = norm_shape(ws)
    ss = norm_shape(ss)
    ws = np.array(ws)
    ss = np.array(ss)
    shape = np.array(a.shape)

    # certificar que ws, ss, and a.shape tem o mesmo numero de dimensoes
    ls = [len(shape),len(ws),len(ss)]
    if 1 != len(set(ls)):
        raise ValueError('a.shape, ws e ss devem ter a mesma dimensao. Eles foram %s' %

    # certificar que ws e menor que a em todas as dimensoes
    if np.any(ws > shape):
        raise ValueError('ws nao pode ser maior que a em qualquer dimensao. a.shape foi
```

```
# pra cada dimensao, criar uma lista de todos os slices validos
slices = [[] for i in range(len(ws))]
for i in xrange(len(ws)):
    nslices = ((shape[i] - ws[i]) // ss[i]) + 1
    for j in xrange(0,nslices):
        start = j * ss[i]
        stop = start + ws[i]
        slices[i].append(slice(start,stop))
# gerar um iterador sobre todos os slices validos
allslices = product(*slices)

# alocar memoria para armazenar todos os slices validos
nslices = np.product([len(s) for s in slices])
out = np.ndarray((nslices,) + tuple(ws),dtype = a.dtype)
for i,s in enumerate(allslices):
    out[i] = a[s]

return out

# In[189]:

#Aplicacao de media movel com janela de 128 observacoes
xf = df.rolling(min_periods=1, center=True, window=128).mean()

#Aplicacao de desvio padrao com janela de 128 observacoes
for i in range(xf.shape[1]):
    xf[i+xf.shape[1]] = df[i].rolling(min_periods=1, center=True, window=128).std()

#Aplicacao da obtencao de valores maximos com janela de 128 observacoes
for i in range(df.shape[1]):
    xf[i+xf.shape[1]] = df[i].rolling(min_periods=1, center=True, window=128).max()

#Aplicacao da obtencao de valores minimos com janela de 128 observacoes
for i in range(df.shape[1]):
    xf[i+xf.shape[1]] = df[i].rolling(min_periods=1, center=True, window=128).min()

#Aplicacao da obtencao da correlacao entre os valores na janela de 128 observacoes
for i in range(df.shape[1]):
    xf[i+xf.shape[1]] = df[i].rolling(min_periods=1, center=True, window=128).corr()

# In[190]:

#Selecao da parcela de 70% dos dados para treinamento do classificador
df_length = xf.shape[0]
train_length = int(df_length*0.7)

x_train = xf[:train_length]
y_train = y[:train_length]

#Selecao da parcela de dados referentes ao teste do classificador
x_test = xf[train_length:]
y_test = y[train_length:]
```

```
# In[191]:

#normalizacao dos valores
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
scaler.fit(x_train)
x_train = scaler.transform(x_train)
x_test = scaler.transform(x_test)

#clf = RandomForestClassifier(n_estimators=40)
#Aplicacao do classificador de arvore de decisao
clf = DecisionTreeClassifier()
#clf = svm.SVC()
#clf = MLPClassifier(solver='lbfgs', alpha=1e-5,hidden_layer_sizes=(500,), random_state=1)

start = time()
clf = clf.fit(x_train, y_train)

print('O Classificador levou %.2f segundos para ser executado' % ((time() - start)))
y_predict = clf.predict(x_test)

print(classification_report(y_test, y_predict))

import itertools
import numpy as np
import matplotlib.pyplot as plt

from sklearn import svm, datasets
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix

class_names = ["caminhando", "movimentando", "parado"]

def plot_confusion_matrix(cm, classes,
                        normalize=False,
                        title='Matriz de Confusão',
                        cmap=plt.cm.Blues):
    """
    Funcao para protagem da matriz de confusao
    A normmalizacao pode ser configurada usando
    normalize=True`.
    """

    plt.imshow(cm, interpolation='nearest', cmap=cmap)
    plt.title(title)
    plt.colorbar()
    tick_marks = np.arange(len(classes))
    plt.xticks(tick_marks, classes, rotation=45)
    plt.yticks(tick_marks, classes)

    if normalize:
```

```
    cm = cm.astype('float') / cm.sum(axis=1)[:, np.newaxis]
    print("Matriz de Confusao Normalizada")
else:
    print('Matriz de Confusao, Sem Normalizacao')

print(cm)

thresh = cm.max() / 2.
for i, j in itertools.product(range(cm.shape[0]), range(cm.shape[1])):
    plt.text(j, i, cm[i, j],
             horizontalalignment="center",
             color="white" if cm[i, j] > thresh else "black")

plt.tight_layout()
plt.ylabel('Atividade Real')
plt.xlabel('Atividade Reconhecida')

# In[193]:

# Computar matriz de confusao
cnf_matrix = confusion_matrix(y_test, y_predict)
np.set_printoptions(precision=2)

# Plotagem da Matriz de Confusao sem Normalizacao

plt.figure(figsize=(10,10))
plot_confusion_matrix(cnf_matrix, classes=class_names,
                      title='Matriz de Confusao, Sem Normalizacao')

# Plotagem da Matriz de Confusao com Normalizacao
plt.figure(figsize=(10,10))
plot_confusion_matrix(cnf_matrix, classes=class_names, normalize=True,
                      title='Matriz de Confusao Normalizada')

plt.show()
```