

CASE – BOLSISTA GRADUADO – VISÃO COMPUTACIONAL - 2025

Reconhecimento e Monitoramento de Posturas Humanas

Nome Completo: Anna Luiza Gomes da Silva

Data: 25/05/2025

1. Introdução

A crescente necessidade de sistemas de monitoramento postural em ambientes de saúde, como quartos hospitalares, impulsiona a criação de modelos robustos de visão computacional. A primeira etapa crucial para o desenvolvimento desses modelos é a seleção e preparação de um conjunto de dados adequado. A presente atividade (AT1) teve como foco a criação de um subconjunto representativo do MPII Human Pose Dataset (<https://www.mpi-inf.mpg.de/departments/computer-vision-and-machine-learning/software-and-datasets/mpii-human-pose-dataset/download>), concentrando-se nas três classes de postura humana definidas no briefing do case: "em pé", "sentado" e "em movimento".

Este relatório descreve detalhadamente a execução da Atividade 1 (AT1) do case para Bolsista em Visão Computacional, que consiste na obtenção e análise de um conjunto de dados. O objetivo principal foi selecionar, organizar e analisar um subconjunto específico do MPII Human Pose Dataset .

O processo envolveu uma exploração inicial do dataset MPII para identificar e mapear as atividades relevantes para as classes de postura desejadas. Posteriormente, um script foi implementado para filtrar e organizar as imagens correspondentes. A etapa final compreendeu uma análise descritiva do conjunto de dados resultante, abordando a distribuição das classes e apresentando exemplos visuais.

O conjunto de dados processado encontra-se agora estruturado e pronto para as próximas fases do projeto, que incluem modelagem e experimentação (AT2).

2. Metodologia de Criação e Organização do Conjunto de Dados

2.1. Fonte de Dados: MPII Human Pose Dataset

O MPII Human Pose Dataset foi selecionado como a fonte de dados para esta atividade. Este dataset é amplamente reconhecido na área de visão computacional e oferece um vasto número de imagens de atividades humanas em cenários reais, acompanhadas de anotações detalhadas, incluindo a atividade realizada pela pessoa em cada imagem. A disponibilidade desses rótulos de atividade foi fundamental para o processo de filtragem e mapeamento para as classes de postura de interesse.

2.2. Definição e Mapeamento das Classes de Postura

Para atender ao requisito do projeto de identificar as posturas "em pé", "sentado" e "em movimento", foi realizada uma análise exploratória inicial (documentada em notebooks/data_analysis.ipynb) para mapear estas categorias às atividades do MPII. Este processo envolveu a inspeção das descrições de atividade (act_name) e a análise visual de imagens do dataset, visando selecionar atividades do MPII que fossem semanticamente correspondentes e visualmente representativas das posturas alvo.

Com base nessa análise, as seguintes atividades presentes no MPII foram mapeadas para as classes alvo:

- **Sentado:**
 - sitting quietly
 - sitting, talking in person, on the phone, computer, or text messaging, light effort
- **Em Pé:**
 - paddle boarding, standing
 - standing, doing work
- **Em Movimento:**
 - walking, for exercise, with ski poles (Observação: a grafia "exerceise" foi mantida conforme o dataset MPII para garantir correspondência exata)
 - skating, ice dancing

Este mapeamento foi implementado no script `code/filter_images.py` através do dicionário `specific_activity_to_class_map`.

2.3. Processamento, Filtragem e Organização das Imagens

O script `code/filter_images.py`, que utiliza a classe `PoseDatasetProcessor`, foi desenvolvido para automatizar a criação do dataset. As principais etapas executadas por este script são:

- **Carregamento das Anotações:** O arquivo `.mat` contendo as anotações do MPII (`mpii_human_pose_v1_u12_1.mat`) é carregado.
- **Extração de Informações:** Para cada anotação, o nome do arquivo da imagem e a string da atividade correspondente são extraídos.
- **Validação de Arquivos:** É verificada a existência do arquivo de imagem correspondente no diretório `data/images/`. As imagens não encontradas são descartadas.
- **Filtragem por Atividade:** As imagens são selecionadas com base no mapeamento definido na seção 2.2. Apenas as imagens cuja atividade corresponde a uma das atividades mapeadas são mantidas.
- **Criação do DataFrame:** As informações das imagens filtradas (nome do arquivo, atividade original e a classe mapeada) são armazenadas em um DataFrame do Pandas.
- **Remoção de Duplicatas:** Caso uma mesma imagem apareça múltiplas vezes nas anotações (por exemplo, com anotações de várias pessoas, mas com a mesma atividade principal), apenas a primeira ocorrência é mantida, garantindo que cada imagem seja associada a uma única classe no dataset final.

- **Criação da Estrutura de Diretórios:** Uma estrutura de diretórios é criada em `data/processed_mpii_poses/` para organizar as imagens:

```
processed_mpii_poses/
├── standing/
│   ├── train/
│   ├── val/
│   └── test/
├── sitting/
│   ├── train/
│   ├── val/
│   └── test/
└── walking/
    ├── train/
    ├── val/
    └── test/
```

- **Divisão em Conjuntos de Treino, Validação e Teste:** Para um desenvolvimento robusto do modelo, o conjunto de dados filtrado foi particionado nas seguintes proporções e finalidades:
 1. Treino (70%): Destinado ao aprendizado do modelo, onde seus parâmetros internos (pesos e vieses) são ajustados iterativamente.
 2. Validação (20%): Utilizado para a otimização de hiperparâmetros (configurações externas que governam o processo de treinamento, como taxa de aprendizado ou arquitetura da rede) e para a detecção precoce de sobreajuste (overfitting), auxiliando na seleção do melhor modelo.
 3. Teste (10%): Reservado para a avaliação final e imparcial da capacidade de generalização do modelo em dados completamente novos, fornecendo uma estimativa de seu desempenho em condições reais de uso.
 - a. A estratificação pela coluna class foi aplicada durante a divisão para assegurar que a distribuição proporcional das classes fosse mantida em todos os subconjuntos.
- **Cópia das Imagens:** As imagens são copiadas dos diretórios originais (data/images/) para as respectivas pastas de classe e divisão dentro de data/processed_mpii_poses/.
- **Geração de Plots:** Plots da distribuição das classes e exemplos de imagens por classe são gerados e salvos em data/processed_mpii_poses/plots/.

3. Análise Descritiva do Conjunto de Dados Criado

Após a execução do script filter_images.py, o conjunto de dados processado foi analisado.

3.1. Distribuição das Classes

Um total de 363 imagens únicas foram selecionadas e processadas. A distribuição dessas imagens entre as classes definidas é apresentada abaixo:

- Standing: 110 imagens
- Sitting: 106 imagens
- Walking: 143 imagens

A visualização gráfica da distribuição das classes é mostrada na Figura 1.

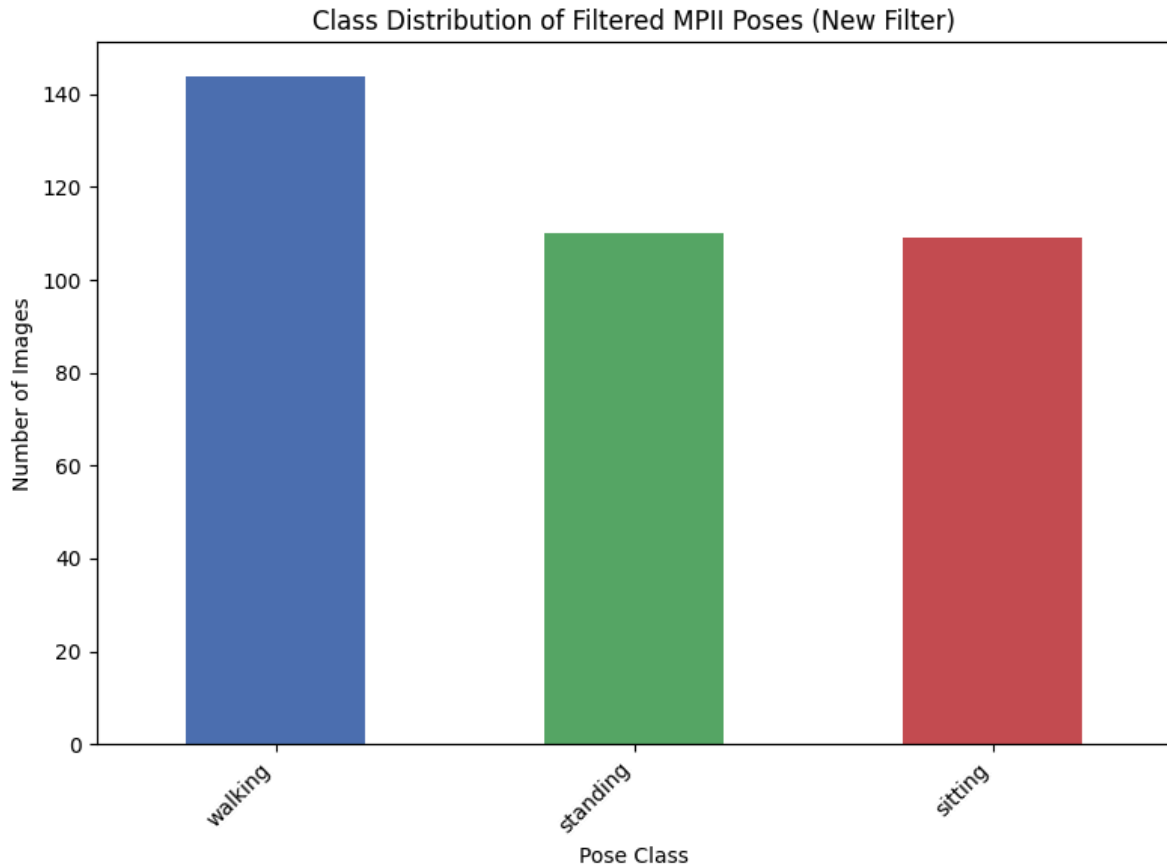


Figura 1: Distribuição das classes no conjunto de dados filtrado.

Comentário sobre Balanceamento:

A análise da distribuição das classes (Figura 1) revela um leve desbalanceamento no conjunto de dados. A classe walking apresenta o maior número de amostras (143 imagens), seguida por standing (110 imagens) e sitting (106 imagens). Embora a diferença entre standing e sitting seja mínima, ambas possuem aproximadamente 23-26% menos amostras que a classe walking. Este desbalanceamento, embora não extremo, deve ser considerado durante a fase de modelagem. Estratégias como a utilização de pesos de classe inversamente proporcionais à frequência das classes no conjunto de treinamento, ou técnicas de reamostragem (oversampling da minoria ou undersampling da maioria), podem ser avaliadas para mitigar potenciais vieses do modelo em favor da classe majoritária.

3.2. Exemplos Visuais por Classe

Para fornecer uma compreensão qualitativa das imagens em cada classe, exemplos visuais são apresentados na Figura 2.

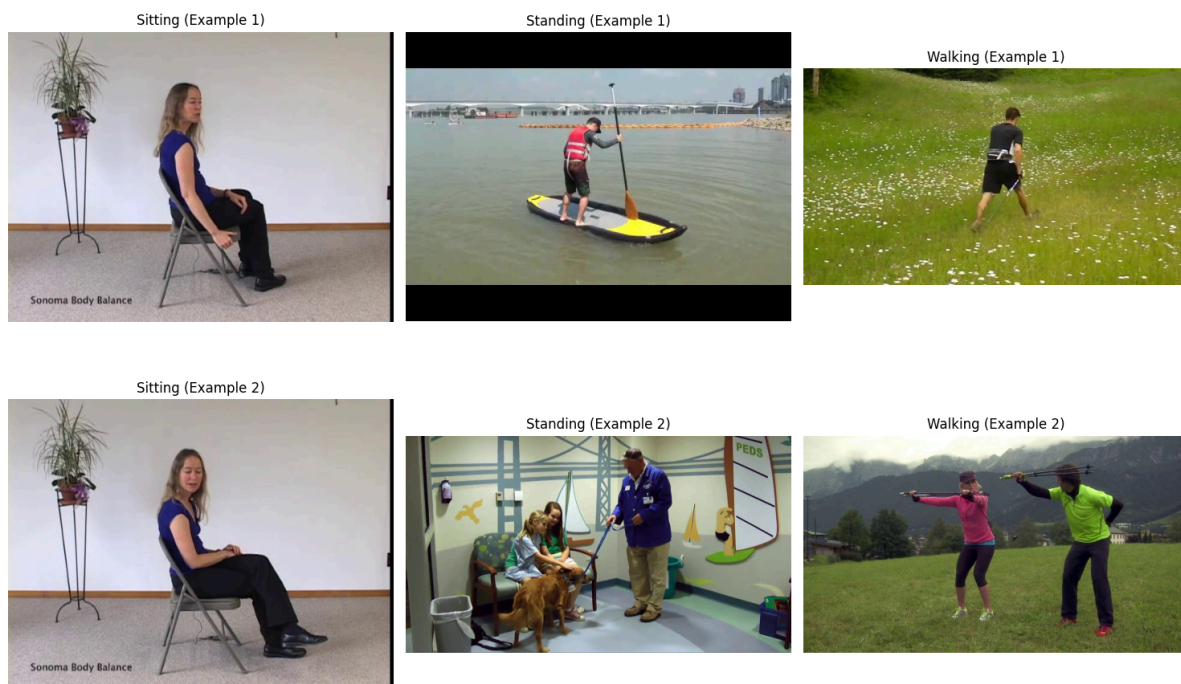


Figura 2: Exemplos de imagens para cada classe de postura.

Comentário sobre Qualidade e Variedade:

Os exemplos visuais (Figura 2) ilustram a diversidade de cenários e contextos presentes no conjunto de dados para cada classe de postura.

Observa-se:

Sitting: Exemplos em ambientes internos controlados, com a postura claramente definida em uma cadeira. A iluminação parece adequada e a pessoa está em foco.

Standing: Um exemplo é em um ambiente externo (paddle boarding), demonstrando a variabilidade de fundos e atividades associadas à postura. O outro exemplo é em um ambiente interno mais complexo (clínica), com múltiplas pessoas e objetos, onde a postura "em pé" do indivíduo principal é identificável, mas com potencial para oclusão parcial ou distração por outros elementos na cena.

Walking: Ambos os exemplos são em ambientes externos com iluminação natural variada. As posturas de movimento são dinâmicas e capturadas em diferentes fases da caminhada/corrida.

De modo geral, as imagens apresentam uma variedade de condições de iluminação (interna, externa/sol, externa/nublado) e complexidade de cena. As posturas, na maioria dos exemplos, são distinguíveis. No entanto, a presença de múltiplos indivíduos (como no exemplo de Standing (Example 2)) e a natureza dinâmica da classe Walking podem introduzir desafios como oclusão parcial, variações de escala e a necessidade de identificar corretamente o sujeito principal da ação. Essa variabilidade é representativa dos desafios

encontrados em dados do mundo real e contribui para a robustez potencial de um modelo treinado com este conjunto.

3.3. Características, Limitações e Ruídos Identificados

Variedade: O MPII Dataset oferece uma boa variedade de cenas internas e externas, diferentes tipos de pessoas e ângulos de câmera, o que contribui para um dataset mais robusto.

Especificidade das Atividades: O mapeamento de atividades específicas do MPII para classes de postura mais genéricas (e.g., "skating, ice dancing" para "walking/em movimento") é uma simplificação. Embora semanticamente razoável para o propósito do case, algumas nuances podem ser perdidas. A classe "em movimento" é inerentemente mais diversa.

Qualidade das Anotações do MPII:

O MPII foca em pose humana, e as anotações de atividade são um metadado adicional. A qualidade e consistência dessas anotações de atividade podem variar.

Representatividade: O dataset filtrado representa um subconjunto do MPII. Embora a estratificação ajude, a representatividade total das posturas em todos os contextos possíveis é limitada pela natureza do dataset original e pelo processo de filtragem.

Multi-Pessoa e Multi-Atividade: O script atual simplifica o tratamento de imagens com múltiplas anotações, mantendo a primeira ocorrência por nome de imagem. Isso implica que se uma imagem tivesse várias pessoas com atividades diferentes, apenas a primeira anotação de atividade (ou a única, se o RELEASE['act'] for por imagem) seria considerada para a classificação da imagem.

4. Estrutura dos Entregáveis (AT1)

Os entregáveis para esta atividade foram organizados da seguinte forma no repositório do projeto:

- Dados originais salvos no google drive:

https://drive.google.com/drive/folders/1ijxSvCOhz_motGe6Dx4b-bVUHKXt1zye?usp=sharing

- Código-fonte da seleção, organização e análise dos dados:

disponível em https://github.com/annalug/AT1_dataset_creation_analysis

code/filter_images.py: Script principal para processamento e organização do dataset.

notebooks/data_analysis.ipynb: Notebook Jupyter utilizado para a exploração inicial dos dados e definição do mapeamento de classes.

- Conjunto de Dados Criado:

Localizado em data/processed_mpii_poses/, contendo as imagens organizadas por classe e divisão (train/val/test) disponíveis em

https://drive.google.com/drive/folders/1ijxSvCOhz_motGe6Dx4b-bVUHKXt1zye?usp=sharing

.

Os plots de análise estão em data/processed_mpii_poses/plots/.

- Relatório em PDF: Este documento.

5. Conclusão e Próximos Passos

A Atividade 1 resultou na criação bem-sucedida de um conjunto de dados estruturado e analisado, contendo imagens das classes de postura "standing", "sitting" e "walking", extraídas e processadas a partir do MPII Human Pose Dataset. A análise descritiva forneceu insights sobre a distribuição das classes e a natureza das imagens.

O conjunto de dados está agora preparado para a Atividade 2 (AT2) – Modelagem e Experimentação, onde diferentes modelos de aprendizado de máquina podem ser treinados e avaliados para a tarefa de reconhecimento de postura humana. As características e limitações identificadas nesta etapa serão consideradas durante o desenvolvimento e avaliação dos modelos.