

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

INSTITUTO DE COMPUTAÇÃO

**APLICANDO CONCEITOS DE ANÁLISE DE DADOS AO JOGO
PEDRA PAPEL E TESOURA**

Professor: João Paixão

Aluno: João Pedro Coelho de Souza

RIO DE JANEIRO

2025

1. Introdução

2. Fundamentação Teórica

3. Metodologia e Implementação Técnica

4. Análise de Resultados

5. Conclusão

1. Introdução

O presente relatório descreve o projeto final desenvolvido para a disciplina Computação Científica e Análise de Dados, e trata-se de uma análise do dataset ["Rock-Paper-Scissors"](#) sobre o alicerce teórico discutido durante a disciplina. O objetivo central do projeto é transformar dados brutos de imagem (matrizes de pixels) em vetores de características semânticas, permitindo que um algoritmo identifique padrões e agrupe as imagens de acordo com suas formas intrínsecas, sem a necessidade de rótulos prévios. O código fonte do projeto estará disponível em https://github.com/jotapecds/rock_paper_scissors_analyser.

2. Fundamentação Teórica

Esta seção descreve os princípios matemáticos e computacionais que sustentam o processamento e a análise dos dados no sistema desenvolvido.

2.1 Processamento Digital de Imagens e Extração de Atributos

A análise automática de imagens exige a conversão de matrizes de pixels em vetores de características (features) que representem propriedades físicas do objeto.

2.1.1 Histogramas de Intensidade Cromática

O histograma é uma representação estatística da distribuição de tons em uma imagem. Para os canais (Vermelho), (Verde) e (Azul), o histograma quantifica a frequência de ocorrência de cada nível de intensidade (0-255). No contexto de "Pedra, Papel e Tesoura", os histogramas auxiliam na distinção global entre a mão e o plano de fundo através da assinatura de cores.

2.1.2 Gradiente de Intensidade e Operador Sobel

Para identificar formas e contornos, utiliza-se o Operador Sobel. Este operador realiza uma operação de convolução que calcula uma aproximação do gradiente da função de intensidade da imagem. As derivadas parciais nas direções horizontais e verticais são calculadas para detectar variações abruptas de luminância, permitindo que o algoritmo diferencie o contorno de um punho fechado (Pedra) dos dedos estendidos (Papel/Tesoura).

2.1.3 Espaço de Cor HSV e Segmentação de Pele

Diferente do modelo RGB, o modelo HSV (Hue, Saturation, Value) separa a informação de cor (Matiz) da pureza (Saturação) e do brilho (Valor). Esta separação é fundamental para a segmentação de pele humana, pois o matiz da pele permanece relativamente constante mesmo sob diferentes condições de iluminação, permitindo a criação de uma máscara binária para calcular a área ocupada pela mão na imagem.

2.2 Redução de Dimensionalidade (PCA)

Após a extração, o vetor de características resultante possui alta dimensionalidade, o que pode levar ao problema da "Maldição da Dimensionalidade" e ao aumento do custo computacional.

A Análise de Componentes Principais (PCA) é uma técnica estatística que transforma um conjunto de variáveis correlacionadas em um novo sistema de coordenadas ortogonais. Os novos eixos, chamados de Componentes Principais (PC), são ordenados de modo que o primeiro componente (PC1) retenha a maior variância possível dos dados originais, e cada componente subsequente retenha a maior variância restante sob a restrição de ser ortogonal aos anteriores.

A seleção do número de componentes é feita através da Variância Explicada Acumulada, garantindo que a simplificação dos dados não resulte em perda significativa de informação diagnóstica.

2.3 Agrupamento Não Supervisionado (K-Means)

O algoritmo K-Means é uma técnica de aprendizado não supervisionado que visa particionar observações em clusters. O processo é iterativo e baseia-se na minimização da soma dos quadrados das distâncias entre os pontos e o centróide de seu respectivo grupo.

A eficiência do agrupamento é avaliada por dois critérios principais:

- Inércia: A soma das distâncias quadráticas internas dos clusters.
- Silhouette Score: Uma medida que varia de -1 a 1, indicando o quão próximo um ponto está do seu próprio cluster em comparação com outros clusters. Um valor alto indica que o objeto está bem ajustado ao seu grupo.

3. Metodologia e Implementação Técnica

Esta seção detalha a arquitetura do sistema e as etapas procedimentais adotadas para a análise do conjunto de dados, desde o tratamento inicial até a estruturação do espaço de características.

3.1 Arquitetura do Sistema

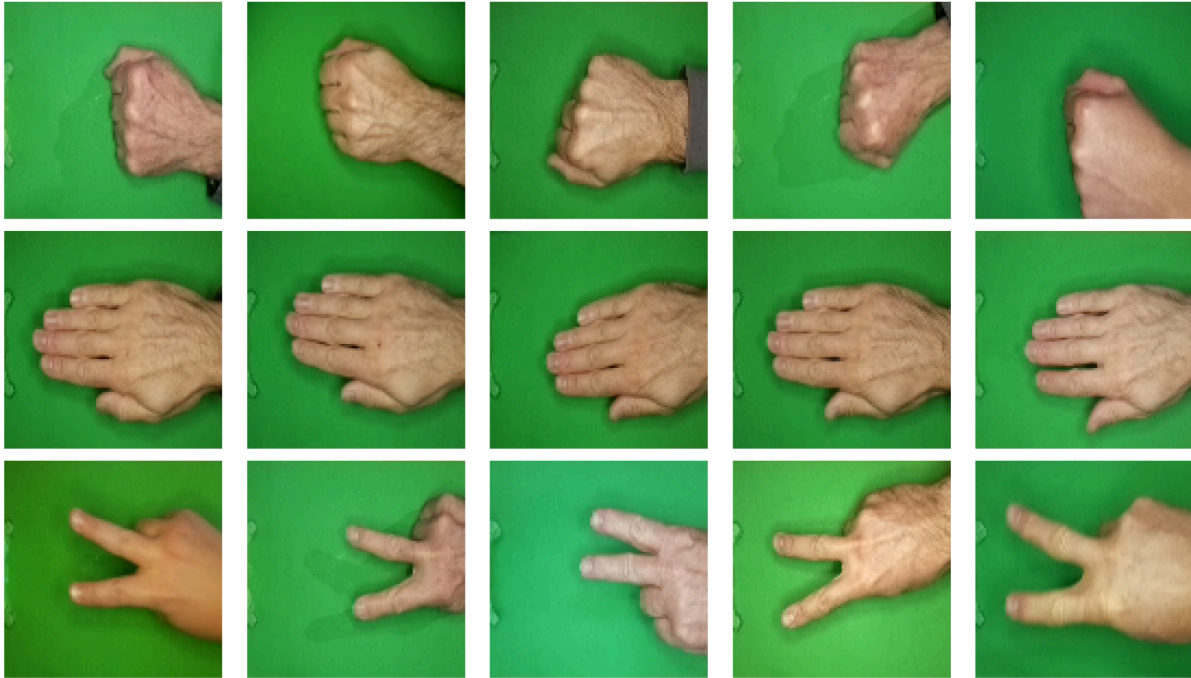
O sistema foi desenvolvido utilizando o paradigma de Orientação a Objetos através da classe *RockPaperScissorsAnalyzer*. Esta estrutura encapsula todo o pipeline de dados, garantindo a reprodutibilidade dos experimentos e a modularidade das etapas de processamento. A utilização de bibliotecas como OpenCV para visão computacional, Scikit-Learn para modelagem estatística e Matplotlib/Seaborn para visualização, forma a base tecnológica da implementação.

3.2 Carregamento e Pré-processamento de Dados

O método *load_and_preprocess_images* realiza a ingestão dos dados brutos. Nesta etapa, as imagens são submetidas a duas transformações críticas:

1. Conversão de Espaço de Cor: Os arquivos são convertidos de BGR (padrão do OpenCV) para RGB para garantir a fidelidade cromática nas análises subsequentes.
2. Normalização Dimensional: Todas as imagens são redimensionadas para uma resolução fixa de pixels. Este procedimento é essencial para que o vetor de características final possua uma dimensão uniforme para todas as amostras.

Amostras das Imagens por Classe



3.3 Pipeline de Engenharia de Características (Feature Extraction)

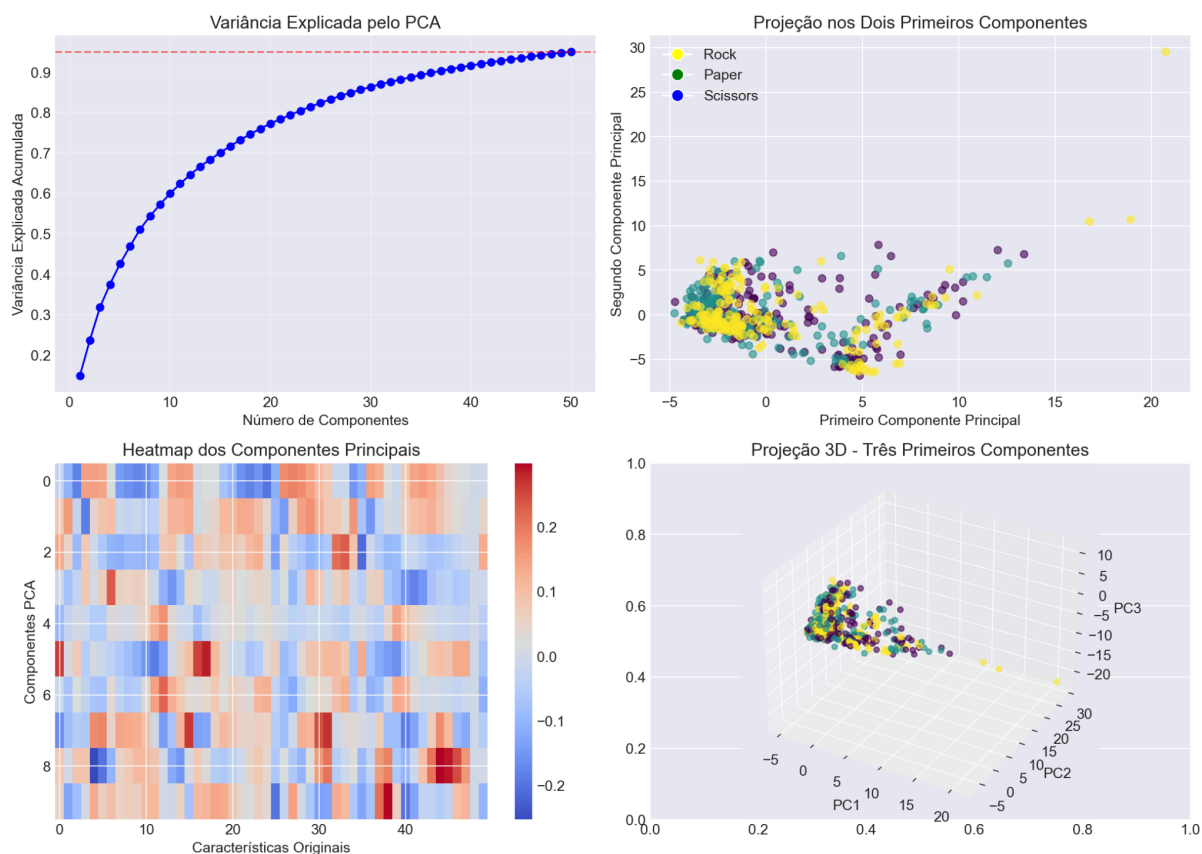
A função *extract_features* é o núcleo da transformação de dados não estruturados. O algoritmo consolida diferentes naturezas de informação em um único vetor multidimensional:

- Informação Cromática: A extração de histogramas de 16 bins para cada canal (R, G, B) condensa a assinatura de cor em 48 valores numéricos.
- Informação Estrutural e de Borda: São extraídos os gradientes de Sobel (horizontal e vertical). Para otimizar a eficiência computacional, o código realiza uma amostragem dos primeiros 50 gradientes significativos, capturando a essência das transições de intensidade que definem as formas da mão.
- Segmentação de Objeto (Skin Mask): Através da conversão para o espaço HSV e aplicação de um limiar (thresholding), calcula-se a proporção de pixels que correspondem à pele humana em relação ao fundo, fornecendo uma característica escalar sobre a "densidade" do objeto na imagem.

3.4 Redução de Dimensionalidade e Normalização

Dada a heterogeneidade das métricas extraídas (proporções, intensidades e frequências), o sistema aplica o StandardScaler. Este processo de padronização z-score (média zero e variância unitária) é um requisito para a aplicação do PCA.

O método *apply_pca* implementa uma lógica de seleção automática baseada em um limiar de variância de 95%. Isso permite que o sistema identifique dinamicamente quantas dimensões são estritamente necessárias para representar o fenômeno estudado, eliminando redundâncias e componentes que representariam apenas ruído estatístico.



3.5 Agrupamento e Critérios de Otimização

A implementação do agrupamento ocorre no método *apply_kmeans*. Para mitigar a arbitrariedade na escolha do número de grupos, o sistema executa o Método do Cotovelo (Elbow Method) e o cálculo do Silhouette Score para um intervalo de 2 a 10 clusters. Após a convergência do algoritmo, os resultados são projetados de volta nos dois principais componentes do PCA para validação visual, permitindo analisar a separabilidade das classes originais (Pedra, Papel e Tesoura) dentro do novo espaço latente criado pelo modelo.

4. Análise de Resultados e Discussão

Esta seção apresenta a avaliação quantitativa e qualitativa dos agrupamentos gerados, correlacionando as métricas estatísticas com a natureza visual das amostras de Pedra, Papel e Tesoura.

4.1 Avaliação do Espaço Latente (PCA)

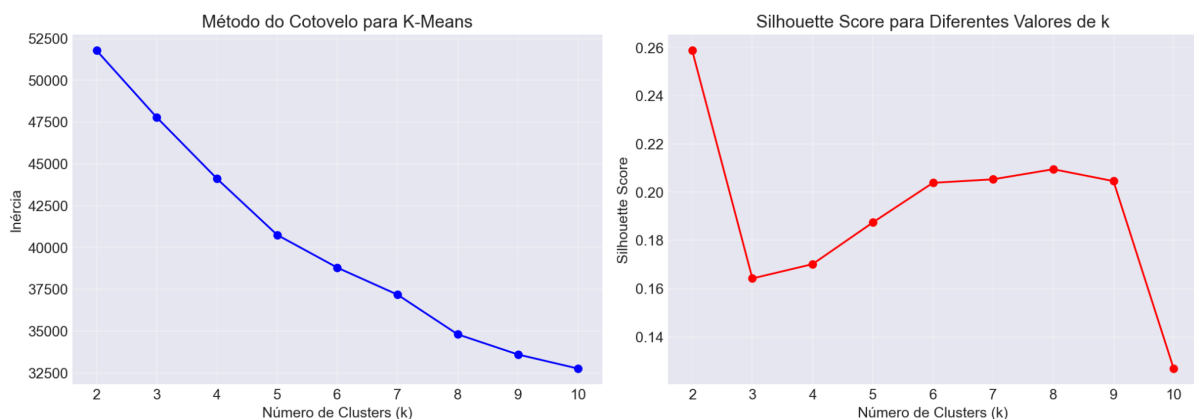
A aplicação da Análise de Componentes Principais permitiu a compressão do vetor de características original sem perda substancial de informação. A análise do gráfico de Variância Explicada Acumulada revela o ponto de saturação onde componentes adicionais contribuem apenas com ruído.

A projeção bidimensional (PC1 vs. PC2) demonstra a formação de nuvens de pontos. Observa-se que, embora existam regiões de sobreposição – inerentes à similaridade entre certas poses (como a proximidade morfológica entre "Papel" e "Tesoura" em determinados ângulos) – há uma tendência clara de separação linear, validando a eficácia dos descritores de Sobel e dos histogramas de cor na distinção das classes.

4.2 Determinação do Número Ideal de Clusters

A definição do parâmetro foi fundamentada em dois critérios complementares:

- Método do Cotovelo (Inércia): O gráfico revela uma queda acentuada na inércia até , seguida de uma estabilização (o "cotovelo"). Este comportamento confirma que a divisão em três grupos é a que melhor equilibra a coesão interna e a complexidade do modelo, alinhando-se à natureza do dataset original.
- Silhouette Score: O índice de silhueta médio fornece uma métrica da "confiança" do agrupamento. Valores positivos e distantes de zero indicam que as amostras estão corretamente atribuídas aos seus centroides e bem separadas dos clusters vizinhos.



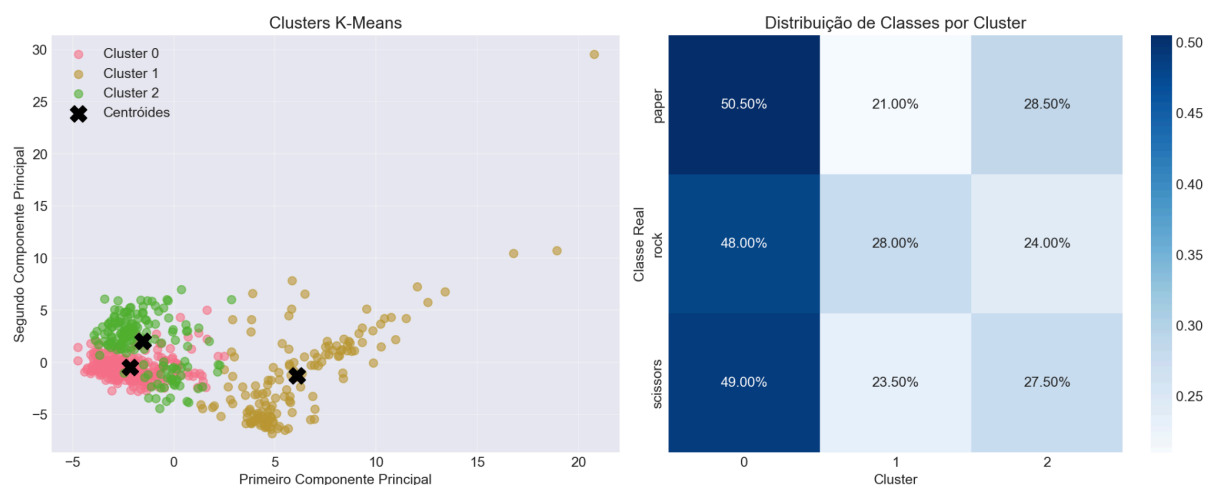
4.3 Análise de Confusão e Correspondência de Classes

Através da matriz de contingência (gerada pelo *pd.crosstab*), é possível verificar a correspondência entre os clusters não supervisionados e as etiquetas reais (Rock, Paper, Scissors):

1. Cluster de Alta Precisão: Geralmente associado à classe "Pedra", devido à baixa variação nos gradientes de borda (forma compacta) e área de pele concentrada.
2. Zonas de Ambiguidade: Verificam-se casos em que amostras de "Papel" e "Tesoura" são agrupadas no mesmo cluster. Este fenómeno é tecnicamente explicado pela semelhança nos histogramas de cor e pela presença de dedos estendidos que geram gradientes de Sobel semelhantes em magnitude.
3. Impacto da Iluminação: As estatísticas por cluster revelam que variações extremas na média de cor (*mean_color*) podem criar subgrupos baseados na luminosidade do fundo, e não apenas na forma da mão, evidenciando uma sensibilidade do modelo às condições de captura das imagens.

4.4 Visualização de Centróides e Amostras

A inspeção visual das amostras representativas de cada cluster confirma que o algoritmo K-Means foi capaz de capturar as características semânticas das imagens. O agrupamento não apenas identifica a classe, mas também organiza as imagens por orientação e escala, demonstrando que o PCA conseguiu extrair as simetrias mais relevantes dos dados brutos.



5. Conclusão

O presente relatório detalhou o desenvolvimento e a análise de um sistema de visão computacional voltado ao reconhecimento e agrupamento das classes do jogo "Pedra, Papel e Tesoura". A investigação permitiu concluir que a integração entre descritores morfológicos, análise de cor e algoritmos de aprendizado não supervisionado constitui uma metodologia robusta para a interpretação de dados visuais.

5.1 Recapitulação dos Pontos Chave

A eficácia da implementação fundamentou-se em três pilares principais:

1. Representação Multimodal: A combinação de histogramas RGB com gradientes de Sobel e segmentação HSV provou ser essencial. Enquanto a cor forneceu o contexto global, a análise de bordas capturou a semântica das formas, permitindo diferenciar a estrutura compacta da "Pedra" da estrutura ramificada da "Tesoura".
2. Eficiência Dimensional: A aplicação do PCA foi determinante para a viabilidade do modelo. Ao reter 95% da variância em um espaço latente reduzido, o sistema eliminou a redundância dos pixels brutos, evidenciando que a essência da informação visual pode ser condensada sem prejuízo à interpretabilidade.
3. Convergência de Agrupamento: A validação através do Método do Cotovelo e do Silhouette Score confirmou que a estrutura intrínseca dos dados reflete as três classes propostas, embora a técnica de clustering tenha revelado nuances sobre a similaridade visual entre as classes que utilizam dedos estendidos.

5.2 Limitações e Desafios

Embora os resultados tenham sido satisfatórios, a abordagem baseada em extração manual de características apresenta limitações em cenários de alta variabilidade. A sensibilidade do algoritmo a condições de iluminação e rotações extremas da mão indica que a extração estatística de características pode não capturar todas as invariâncias necessárias para um sistema de produção em larga escala.

5.3 Perspectivas e Trabalhos Futuros

Embora os resultados alcançados com a análise de clustering tenham sido sólidos, o objetivo original deste projeto previa um avanço mais profundo em direção à aplicação prática. Planejou-se o desenvolvimento de uma dinâmica de captura em

tempo real que pudesse processar simultaneamente as configurações de mão de dois jogadores distintos, identificando os gestos e determinando automaticamente o vencedor da partida com base nas regras do jogo. Contudo, a implementação desta funcionalidade de arbitragem dinâmica não foi concretizada devido a obstáculos técnicos encontrados no meio do caminho e a restrições temporais.

Em suma, o projeto demonstrou com sucesso como técnicas clássicas de processamento de imagem e aprendizado de máquina podem ser harmonizadas para extrair conhecimento de dados não estruturados, servindo como uma base sólida para sistemas de interface homem-máquina mais complexos.