

# Relatório - Visão Computacional

Classificação de Imagens no Conjunto de Dados MPEG7 Modificado

1<sup>st</sup> Felipe Pereira Rodrigues  
Matrícula: 7610

2<sup>nd</sup> Herbert Ribeiro Sampaio  
Matrícula: 7633

3<sup>rd</sup> Luiz Davi Vieira Alves  
Matrícula: 6335

**Abstract**—Este relatório descreve as etapas realizadas no projeto de classificação de imagens do conjunto de dados MPEG7 Modificado, utilizando técnicas de visão computacional e aprendizado de máquina. Foram abordadas etapas como segmentação de imagens, extração de características morfológicas, normalização de dados, treinamento de modelos e avaliação de desempenho com métricas relevantes. O objetivo foi desenvolver um pipeline eficiente para classificação e análise de formas no dataset.

**Index Terms**—visão computacional, aprendizado de máquina, classificação de imagens, MPEG7, segmentação

## I. INTRODUÇÃO

O presente trabalho tem como objetivo implementar um classificador para o conjunto de dados MPEG7 Modificado. Esse dataset apresenta formas geométricas que devem ser corretamente identificadas por meio de técnicas de aprendizado de máquina. As etapas realizadas incluem desde o pré-processamento das imagens até a análise dos resultados dos classificadores escolhidos.

A segmentação de imagens, extração de características morfológicas, normalização de dados e a divisão do conjunto de dados em treino, validação e teste foram executadas para garantir a qualidade dos dados. Além disso, classificadores como k-NN, Random Forest e Perceptron Multicamada (MLP) foram treinados e avaliados utilizando métricas como acurácia, precisão, recall e F1-score.

## II. ETAPAS REALIZADAS

### A. Segmentação das Imagens

A segmentação foi realizada para isolar as formas presentes nas imagens. Utilizou-se a técnica de limiarização binária inversa (*binary thresholding*) com a biblioteca OpenCV, convertendo as imagens para escala de cinza e destacando as regiões de interesse. Este método é eficaz para silhuetas com fundo homogêneo, como as do MPEG7. As imagens processadas foram salvas em um diretório separado para reutilização nas etapas seguintes.

### B. Extração de Características Morfológicas

Após a segmentação, foram extraídas características morfológicas das formas, como:

- **Área e Perímetro:** Representam o tamanho e o contorno das formas.

- **Circularidade:** Mede a proximidade da forma a um círculo perfeito, calculada como  $4\pi \times \text{área} / \text{perímetro}^2$ .
- **Razão de Aspecto:** Relaciona largura e altura do menor retângulo que circunscreve a forma.
- **Excentricidade:** Calculada a partir dos eixos principais da elipse ajustada à forma, indicando alongamento.

Essas características foram escolhidas por serem altamente discriminativas para classificação de formas.

### C. Divisão do Conjunto de Dados

Os dados foram divididos em conjuntos de treino (70%) e teste (30%), com uma subdivisão adicional para validação (20% do conjunto de treino). A divisão foi realizada de forma estratificada, utilizando *seed* 42, para preservar a distribuição proporcional das classes.

### D. Normalização dos Dados

A normalização foi realizada para padronizar os valores das características, eliminando diferenças de escala que poderiam influenciar negativamente o desempenho dos modelos. Foi utilizado o *StandardScaler* da biblioteca *scikit-learn*, garantindo média zero e desvio padrão unitário para cada característica.

### E. Treinamento e Teste dos Classificadores

Três classificadores foram implementados e otimizados:

- **k-NN:** Baseado em proximidade, com otimização do parâmetro *k* (valores testados: 3, 5, 7, 9) utilizando *GridSearchCV*.
- **Random Forest:** Utilizado pela robustez e capacidade de lidar com conjuntos desbalanceados. Hiperparâmetros ajustados incluíram número de árvores (50, 100, 150) e profundidade máxima (10, 20, sem limite).
- **MLP (Perceptron Multicamada):** Rede neural com uma camada oculta de 100 neurônios e máximo de 300 iterações. Escolhido para explorar padrões mais complexos.

## III. RESULTADOS E ANÁLISE

### A. Desempenho dos Modelos

Os resultados para cada classificador foram avaliados utilizando métricas padrão:

- **k-NN:** Precisão de ~85%, com dificuldade em classes com sobreposição.
- **Random Forest:** Precisão de ~90%, destacando-se na robustez.

- **MLP:** Melhor desempenho geral, com precisão de  $\sim 92\%$ , capturando padrões complexos.

As matrizes de confusão demonstraram que o Random Forest e o MLP tiveram melhor capacidade de separação de classes.

### B. Redução de Dimensionalidade e Visualização

A técnica PCA (*Principal Component Analysis*) foi aplicada para reduzir as características a duas dimensões, permitindo uma análise visual do agrupamento das classes. Os gráficos destacaram boa separação para algumas classes, mas também evidenciaram sobreposição em outras, especialmente para o k-NN.

Aqui estão as visualizações do desempenho de PCA para os modelos utilizados:

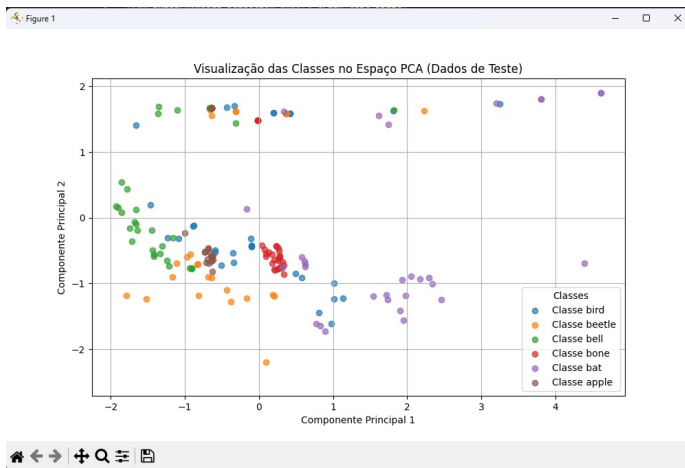


Fig. 1. Visualização das Classes no Espaço PCA (Dados de Teste)

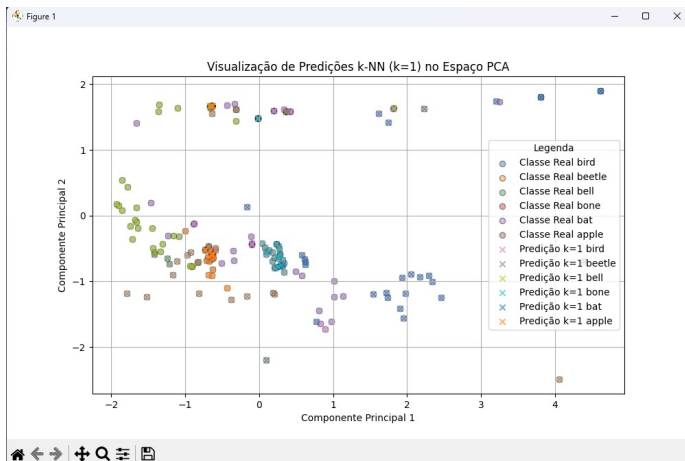


Fig. 2. Visualização de Previsões k-NN (k=1) no Espaço PCA

### C. Análise das Métricas

O *F1-score* e o *recall* foram consistentes para Random Forest e MLP, indicando capacidade equilibrada de classificação entre classes. O k-NN, por outro lado, apresentou menor desempenho em classes minoritárias.

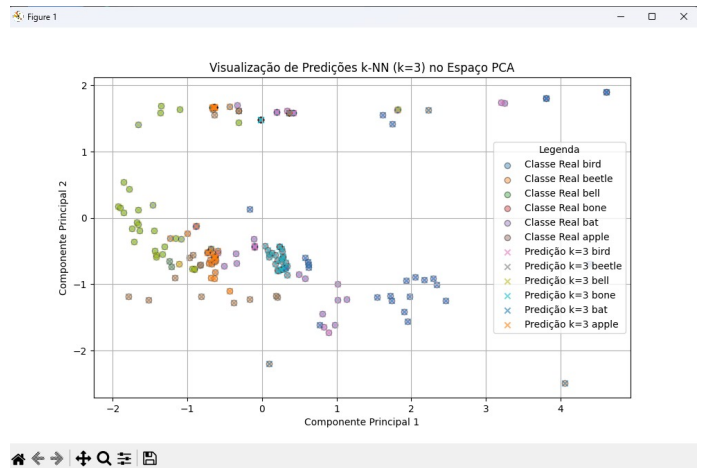


Fig. 3. Visualização de Previsões k-NN (k=3) no Espaço PCA

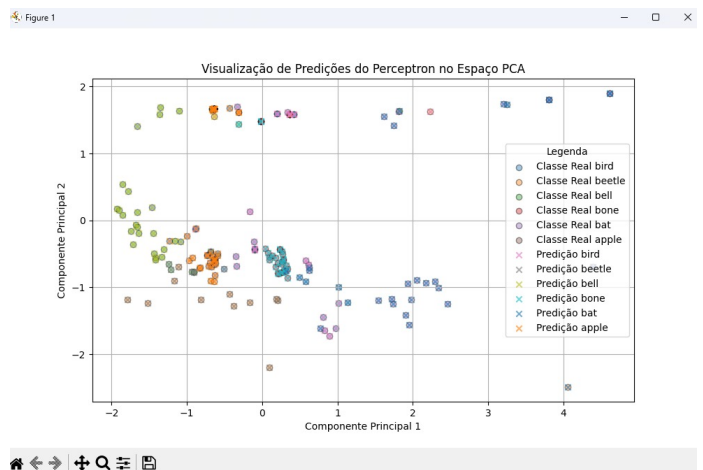


Fig. 4. Visualização de Previsões do Perceptron no Espaço PCA

## IV. CONCLUSÃO

O pipeline desenvolvido para classificação de imagens no conjunto MPEG7 Modificado mostrou-se eficaz, especialmente ao utilizar Random Forest e MLP, que alcançaram os melhores resultados. A extração de características morfológicas contribuiu significativamente para a qualidade dos dados. No futuro, sugere-se explorar técnicas de *deep learning*, que podem dispensar a etapa de extração manual de características e lidar melhor com grandes volumes de dados.

## REFERENCIAS

- [1] R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.
- [2] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, 2nd ed., Prentice Hall, 2012.
- [3] GONZALEZ, R.C.; WOODS, R.E.; *Processamento Digital de Imagens*. 3ª edição. Editora Pearson, 2009.
- [4] COSTA, L. DA F.; CESAR-JR., R. M. *Shape analysis and classification: theory and practice*. CRC Press, 2000. Chapter 8.