

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS

Campus Escola Politécnica
Engenharia de Computação

Igor Parreira de Paula

Reconhecimento de Faces com PCA, SVM e t-SNE

Professor: Clarimar J. Coelho

Goiânia

2025

Igor Parreira de Paula

Reconhecimento de Faces com PCA, SVM e t-SNE

Relatório apresentado como requisito parcial para avaliação na disciplina de Inteligência Artificial (CMP1110), do curso de Engenharia de Computação da PUC Goiás, sob orientação do Prof. Clarimar J. Coelho.

Goiânia

2025

Sumário

1 INTRODUÇÃO	2
1.1 PROBLEMA DE PESQUISA	2
1.2 OBJETIVOS	2
1.2.1 OBJETIVO GERAL	2
1.2.2 OBJETIVOS ESPECÍFICOS	2
2 METODOLOGIA	2
2.1 REFERENCIAL TEÓRICO	2
2.2 PIPELINE DE IMPLEMENTAÇÃO	3
3 RESULTADOS E ANÁLISES	3
3.1 AVALIAÇÃO GERAL E MATRIZ DE CONFUSÃO	3
3.2 ANÁLISE DE ERROS	3
3.3 VISUALIZAÇÃO t-SNE	4
3.4 TESTES DE CLASSIFICAÇÃO DINÂMICA	4
3.4.1 Teste 1: Sujeito s1 (Acerto Esperado)	4
3.4.2 Teste 2: Sujeito s5 (Erro da CV)	6
3.4.3 Teste 3: Sujeito s31 (Aleatório)	6
4 CONCLUSÃO	7

1 INTRODUÇÃO

O reconhecimento facial é uma tarefa fundamental em diversas aplicações de segurança e identificação biométrica. A base de dados ORL (Olivetti Research Laboratory) oferece um conjunto de dados clássico e robusto para experimentações com métodos de visão computacional.

Este trabalho se propõe a desenvolver um sistema de reconhecimento facial aplicando uma combinação de técnicas de Aprendizado de Máquina para classificação eficiente e redução de dimensionalidade.

1.1 PROBLEMA DE PESQUISA

O problema consiste em, dada uma imagem de uma face (um vetor de alta dimensionalidade de $112 \times 92 = 10304$ pixels), identificar corretamente a qual dos 40 sujeitos da base ORL ela pertence. Isso exige um método que reduza a dimensionalidade dos dados (para evitar a "maldição da dimensionalidade") e um classificador robusto que possa lidar com as variações de expressão, iluminação e acessórios (óculos) presentes na base.

1.2 OBJETIVOS

Introduz de modo abrangente os objetivos da pesquisa.

1.2.1 OBJETIVO GERAL

Desenvolver um sistema de reconhecimento facial utilizando a combinação de Análise de Componentes Principais (PCA) com Máquinas de Vetores de Suporte (SVM), aplicadas à base de dados ORL.

1.2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos que guiaram a implementação foram:

- Implementar a leitura e pré-processamento das 400 imagens da base ORL.
- Aplicar redimensionamento para 112×92 pixels em escala de cinza.
- Utilizar PCA para reduzir a dimensionalidade para 50 componentes principais.
- Treinar um classificador SVM (kernel RBF) com os vetores PCA.
- Implementar a classificação de imagens de teste, exibindo o Top-5 de classes prováveis.
- Avaliar o modelo com Validação Cruzada (5-fold) e gerar a Matriz de Confusão.
- Visualizar a separabilidade dos dados com a projeção t-SNE.

2 METODOLOGIA

2.1 REFERENCIAL TEÓRICO

O pipeline de classification baseia-se em duas técnicas centrais:

- **Análise de Componentes Principais (PCA):** Uma técnica de redução de dimensionalidade linear que transforma os dados para um novo espaço de coordenadas, onde os eixos (componentes principais) capturam a máxima variância dos dados. No contexto de faces, isso é conhecido como *Eigenfaces*.

- **Máquinas de Vetores de Suporte (SVM):** Um classificador supervisionado que encontra um hiperplano ótimo para separar classes no espaço de características. O uso do *kernel RBF* (Radial Basis Function) permite ao SVM encontrar fronteiras de decisão não-lineares, o que é ideal para dados complexos como os componentes de faces.

2.2 PIPELINE DE IMPLEMENTAÇÃO

O sistema foi implementado em Python utilizando ‘scikit-learn’ e ‘opencv’. O pipeline seguiu os seguintes passos:

1. **Carga e Pré-processamento:** As 400 imagens foram lidas em escala de cinza, redimensionadas para 112×92 e achataadas em vetores de 10304 dimensões.
2. **Construção do Pipeline:** Foi criado um ‘Pipeline’ do ‘scikit-learn’ contendo:
 - (a) Um estimador ‘PCA(*n_components*=50, whiten=True)’.
 - (b) Um classificador ‘SVC(*kernel*=‘rbf’, probability=True)’.
3. **Treinamento e Avaliação:** O pipeline foi treinado e avaliado usando ‘cross_val_predict’ com ‘KFold(*n_splits*=5)’ para gerar a matriz de confusão e a acurácia geral. Um modelo final também foi treinado com todos os dados para os testes dinâmicos.

3 RESULTADOS E ANÁLISES

O script foi executado e produziu os resultados detalhados a seguir.

3.1 AVALIAÇÃO GERAL E MATRIZ DE CONFUSÃO

A avaliação do modelo utilizando Validação Cruzada de 5 folds (5-fold CV) demonstrou alta performance. A acurácia média obtida foi de **95.75%**.

A matriz de confusão (Figura 1) ilustra visualmente o desempenho. A diagonal principal concentra a vasta maioria das classificações, indicando um alto índice de acerto. Os poucos pontos fora da diagonal representam os erros de classificação.

3.2 ANÁLISE DE ERROS

A saída do console detalha os 17 erros (de 400 imagens) ocorridos durante a validação cruzada.

```
--- Análise de Erros Frequentes (Valida o Cruzada) ---
Total de erros na CV: 17 / 400
Maiores confusões (Real -> Previsto): Freqüência
  s40 -> s5: 4 vez(es)
  s13 -> s1: 2 vez(es)
  s5 -> s40: 1 vez(es)
  s16 -> s1: 1 vez(es)
  s19 -> s40: 1 vez(es)
  s26 -> s28: 1 vez(es)
  s28 -> s26: 1 vez(es)
  s28 -> s37: 1 vez(es)
  s28 -> s17: 1 vez(es)
  s31 -> s38: 1 vez(es)
```

A confusão mais significativa (4 ocorrências) foi o modelo classificar erroneamente imagens do sujeito ‘s40’ como sendo do sujeito ‘s5’. Também houve confusões notáveis entre ‘s13’ e ‘s1’, e ‘s26’ e ‘s28’. Isso sugere uma alta similaridade visual entre esses pares de sujeitos no espaço de características do PCA.

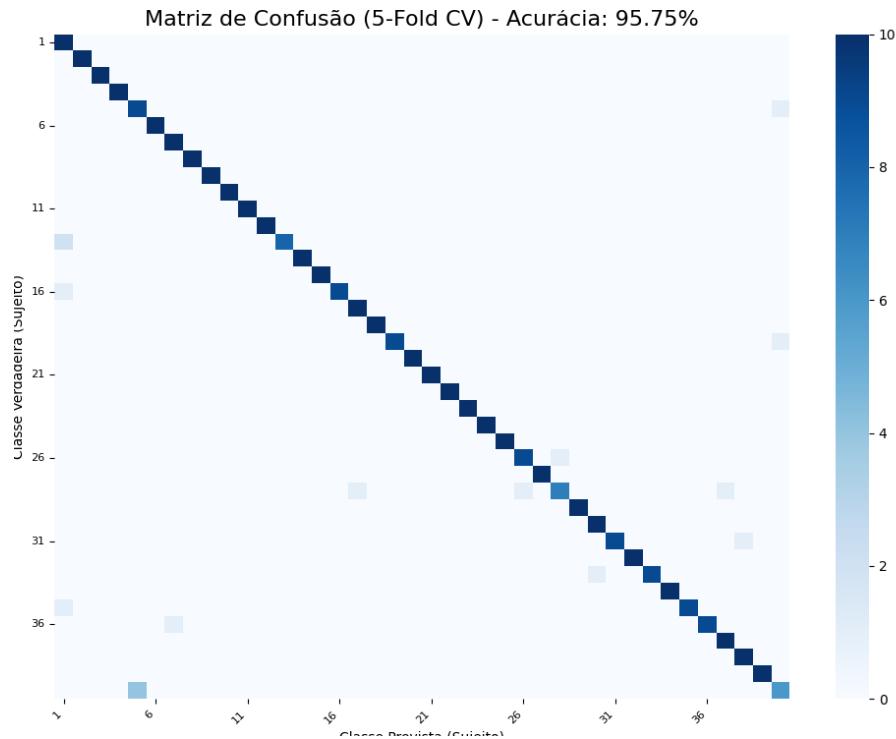


Figura 1: Matriz de Confusão (5-Fold CV) - Acurácia: 95.75%.

3.3 VISUALIZAÇÃO t-SNE

A projeção t-SNE (Figura 2) reduz os 50 componentes do PCA para 2 dimensões, permitindo visualizar a separabilidade das 40 classes.

A visualização mostra que a maioria dos sujeitos forma *clusters* (agrupamentos) coesos e bem separados, explicando a alta acurácia do SVM. No entanto, alguns *clusters* de cores diferentes (sujeitos diferentes) aparecem muito próximos ou sobrepostos, correspondendo às confusões identificadas na seção anterior.

3.4 TESTES DE CLASSIFICAÇÃO DINÂMICA

Foram realizados três testes com imagens específicas.

3.4.1 Teste 1: Sujeito s1 (Acerto Esperado)

Uma imagem do sujeito ‘s1’ foi fornecida ao modelo. A classificação foi **CORRETA**.

```
--- Análise da Imagem de Teste (Sujeito Real: s1) ---
Previsão do Modelo: s1
Resultado: CORRETO
-----
Top-5 Classes Mais Prováveis:
1. Sujeito s1 (Confiança: 52.52%)
2. Sujeito s16 (Confiança: 2.44%)
3. Sujeito s19 (Confiança: 2.29%)
4. Sujeito s24 (Confiança: 2.22%)
5. Sujeito s4 (Confiança: 2.12%)
```

O modelo acertou com 52.52% de confiança, como visto na Figura 3.

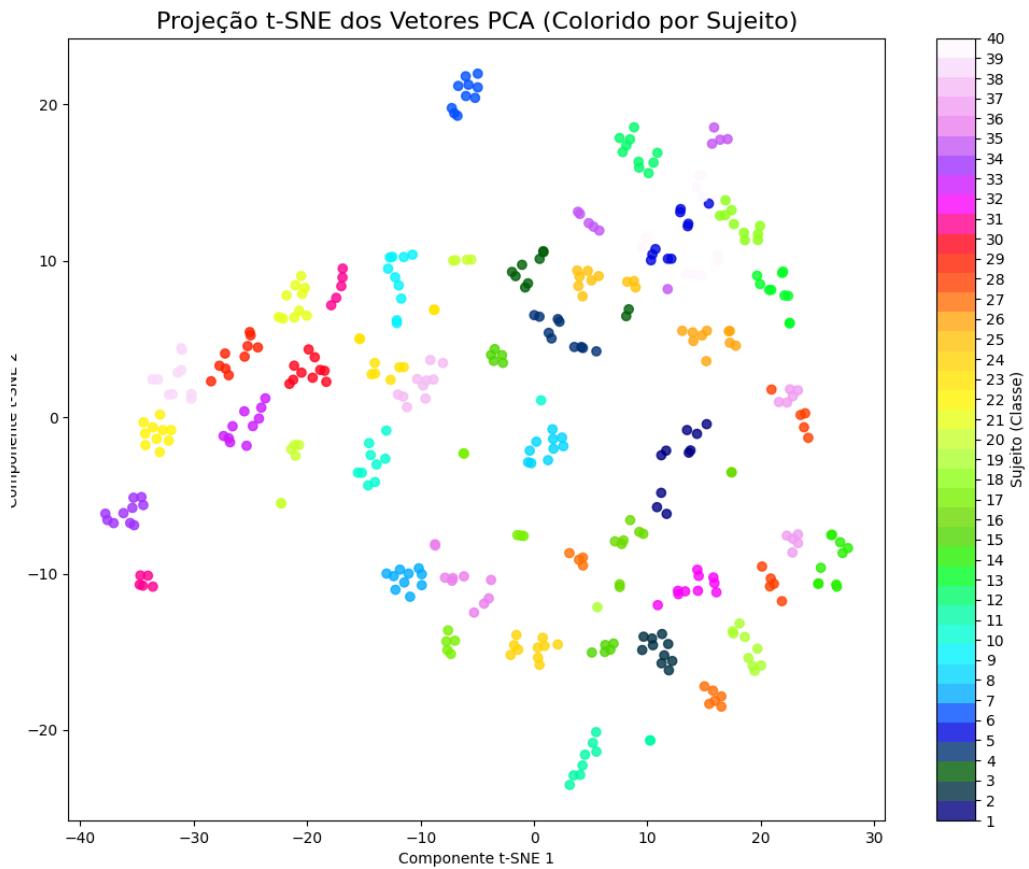


Figura 2: Projeção t-SNE dos 50 componentes PCA coloridos por classe (sujeito).

Análise da Imagem de Teste - Real: s1 | Previsto: s1



Figura 3: Teste 1: Imagem de teste (s1) e as 9 imagens da classe prevista (s1).

3.4.2 Teste 2: Sujeito s5 (Erro da CV)

Uma imagem do sujeito ‘s5’, que foi identificado como um dos sujeitos com erros na CV, foi testada. A classificação foi **CORRETA**.

```
--- Análise da Imagem de Teste (Sujeito Real: s5) ---
Previsão do Modelo: s5
Resultado: CORRETO
-----
Top-5 Classes Mais Prováveis:
1. Sujeito s5 (Confiança: 34.20%)
2. Sujeito s40 (Confiança: 7.75%)
3. Sujeito s18 (Confiança: 3.21%)
4. Sujeito s36 (Confiança: 3.18%)
5. Sujeito s16 (Confiança: 2.37%)
```

Embora o modelo tenha acertado (Figura 4), é notável que a confiança foi mais baixa (34.20%) e o sujeito ‘s40’ (o mais confundido com ‘s5’) apareceu como a segunda classe mais provável (7.75%), validando a análise de erros.

Análise da Imagem de Teste - Real: s5 | Previsto: s5

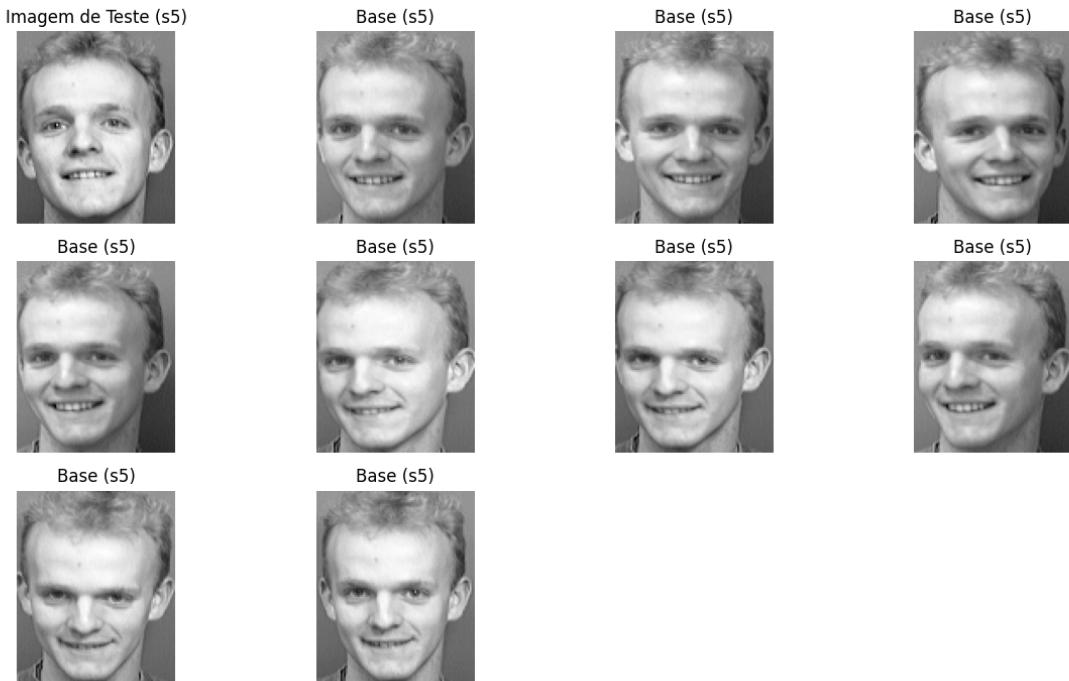


Figura 4: Teste 2: Imagem de teste (s5) e as 9 imagens da classe prevista (s5).

3.4.3 Teste 3: Sujeito s31 (Aleatório)

Uma imagem aleatória do sujeito ‘s31’ foi testada. A classificação foi **CORRETA**.

```
--- Análise da Imagem de Teste (Sujeito Real: s31) ---
Previsão do Modelo: s31
Resultado: CORRETO
-----
Top-5 Classes Mais Prováveis:
1. Sujeito s31 (Confiança: 34.51%)
2. Sujeito s10 (Confiança: 2.57%)
3. Sujeito s26 (Confiança: 2.55%)
```

4. Sujeito s30 (Confiança: 2.49%)
 5. Sujeito s38 (Confiança: 2.49%)

O modelo acertou com 34.51% de confiança (Figura 5).

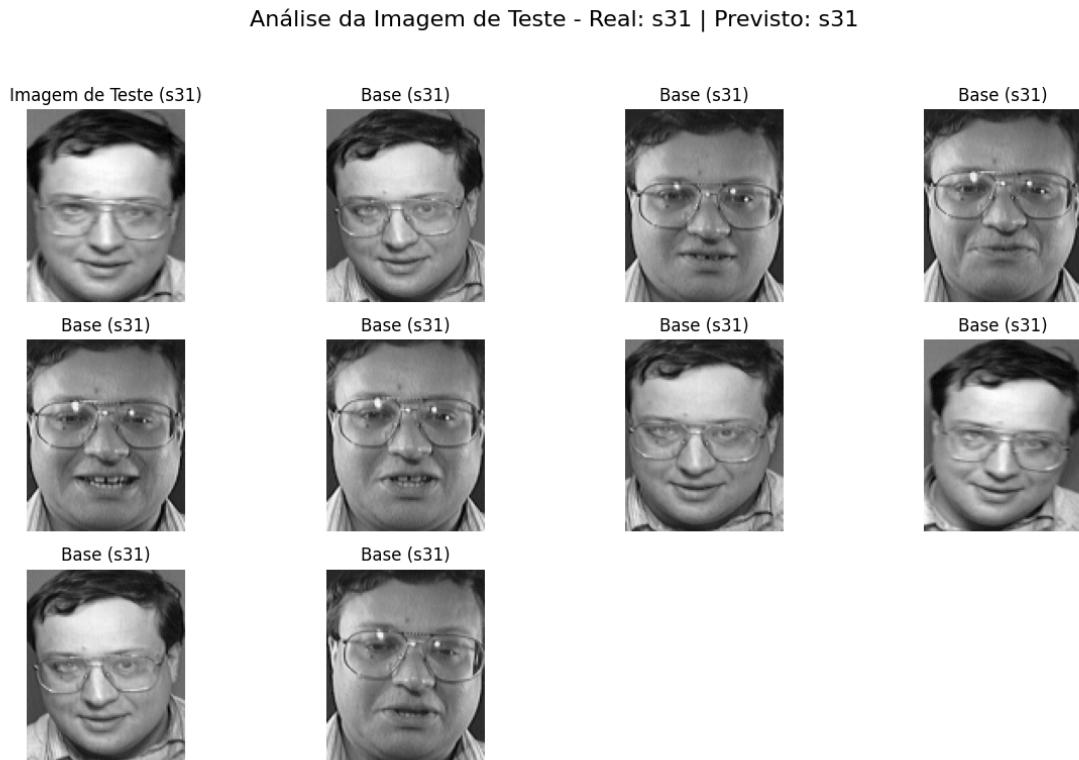


Figura 5: Teste 3: Imagem de teste (s31) e as 9 imagens da classe prevista (s31).

4 CONCLUSÃO

O pipeline de reconhecimento facial utilizando PCA para redução de dimensionalidade e SVM (RBF) para classificação mostrou-se extremamente eficaz para a base de dados ORL, atingindo uma acurácia de 95.75% na validação cruzada.

As análises de erro e a visualização t-SNE confirmam que o modelo aprendeu representações separáveis para a maioria dos sujeitos. As limitações observadas (17 erros) estão concentradas em pares específicos de sujeitos (como ‘s40’ e ‘s5’) que possuem alta similaridade visual, mesmo no espaço reduzido do PCA.

Os testes dinâmicos, embora todos corretos, mostraram diferentes níveis de confiança, refletindo a dificuldade da classificação. O modelo cumpriu todos os requisitos propostos, demonstrando a viabilidade da combinação PCA+SVM para esta tarefa.