



CENTRO UNIVERSITÁRIO DE BRASÍLIA - CEUB

PROGRAMA DE INICIAÇÃO CIENTÍFICA

GABRIEL DE ARAUJO PREDIGER

JOÃO LUCAS GONÇALVES ATAIDE

**COMPARAÇÃO E IMPLEMENTAÇÃO DE DIFERENTES ALGORITMOS DE
RECONHECIMENTO FACIAL**

BRASÍLIA

2023



GABRIEL DE ARAUJO PREDIGER

JOÃO LUCAS GONÇALVES ATAIDE

**COMPARAÇÃO E IMPLEMENTAÇÃO DE DIFERENTES ALGORITMOS DE
RECONHECIMENTO FACIAL**

Relatório final de pesquisa de Iniciação Científica apresentado à Assessoria de Pós-Graduação e Pesquisa.

Orientação: Salvador Alves de Melo Junior

BRASÍLIA

2023

RESUMO

Este artigo tem como objetivo comparar e implementar diferentes algoritmos de reconhecimento facial. Os algoritmos analisados são o Linear Discriminant Analysis (LDA), K Nearest Neighbors (KNN) e Decision Tree Classifier (DTC). O estudo utiliza a base de imagens Yale Faces, composta por 165 imagens de 15 indivíduos em 11 categorias diferentes. O processo de implementação envolveu a padronização das imagens, extração dos rostos e redimensionamento das imagens. Os resultados foram compilados e a acurácia dos algoritmos foi avaliada em diferentes categorias, como luz centrada, óculos, felicidade e luz à esquerda. Conclui-se que o KNN apresentou uma acurácia ligeiramente melhor que o LDA na categoria de luz à esquerda, enquanto o LDA teve a pior acurácia em todas as categorias. Recomenda-se o uso do KNN para casos em que a iluminação esteja mais à esquerda, enquanto o LDA e o KNN podem ser aplicados com boa acurácia em situações com imagens padronizadas e iluminação central, pessoas usando óculos e expressões felizes.

Palavras-chave: reconhecimento facial; LDA (Linear Discriminant Analysis); KNN (K Nearest Neighbors).

LISTAS DE FIGURAS

Figura 1: aplicação do LDA em um conjunto de dados por meio de uma representação visual.....	7
Figura 2: gráfico bidimensional com os eixos X e Y, reduzido a uma dimensão.	8
Figura 3: aplicação alternativa do LDA	8
Figura 4: Variação do Scatter.	9
Figura 5: ponto em azul que será categorizado pelo KNN.....	9
Figura 6: abordagem de regressão do KNN.....	10
Figura 7: abordagem de classificação do KNN.....	10

SUMÁRIO

1. INTRODUÇÃO.....	6
OBJETIVOS	6
2. FUNDAMENTAÇÃO TEÓRICA	7
3. MÉTODO	11
4. RESULTADOS E DISCUSSÃO	13
5. CONSIDERAÇÕES FINAIS (OU CONCLUSÕES).....	14
REFERÊNCIAS	15

1. INTRODUÇÃO

O reconhecimento facial, uma área da computação, tem como objetivo identificar rostos humanos ou pessoas específicas por meio dos padrões presentes nos relevos faciais [1]. Para alcançar esse reconhecimento, diversos métodos e algoritmos podem ser utilizados, como o Linear Discriminant Analysis (LDA) [2,3,4], K Nearest Neighbors (KNN) [5,6,7] e Decision Tree Classifier (DTC) [8,9].

Nesta pesquisa, realizamos uma comparação entre esses algoritmos para determinar qual deles é mais adequado em situações problemáticas específicas encontradas em diversas aplicações de reconhecimento facial. Para isso, seguiremos uma série de passos e etapas, incluindo a busca por exemplos desses algoritmos na Internet, a seleção de uma base de imagens contendo rostos com e sem irregularidades, a extração e padronização das imagens, a aplicação dos algoritmos e a compilação dos resultados obtidos. Com base nessas análises, seremos capazes de inferir qual algoritmo é mais indicado para lidar com as diversas irregularidades encontradas.

Essa abordagem permitirá uma melhor compreensão das capacidades e limitações de cada algoritmo de reconhecimento facial, fornecendo insights valiosos para a seleção e implementação adequada em diferentes cenários de aplicação.

OBJETIVOS

O objetivo desta pesquisa é comparar e implementar diferentes algoritmos de reconhecimento facial, incluindo Linear Discriminant Analysis (LDA), K Nearest Neighbors (KNN) e Decision Tree Classifier (DTC), em diversas situações problemáticas. O estudo busca avaliar o desempenho desses algoritmos em diferentes irregularidades encontradas em aplicações de reconhecimento facial. O artigo tem como meta analisar e comparar os resultados obtidos pelos algoritmos aplicados a uma base de imagens com rostos que possuem e não possuem irregularidades. A partir desses resultados, o objetivo é inferir qual algoritmo é mais adequado para cada tipo de irregularidade encontrada.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta seção, apresentamos os algoritmos escolhidos para o reconhecimento facial, começando pelo Linear Discriminant Analysis (LDA). O LDA é projetado para minimizar as diferenças dentro de uma mesma classe e maximizar as diferenças entre classes distintas. Para alcançar esse objetivo, o LDA identifica o ponto de maior separação entre os dados e cria um novo eixo com dimensão reduzida. Por meio de critérios como a distância entre médias das classes e a minimização da variação interna, o LDA busca obter uma separação ideal. Nesta seção, iremos explorar em detalhes o funcionamento do LDA, incluindo sua aplicação visual em um conjunto de dados e a fórmula utilizada para descrever seus critérios de separação [4].

Linear Discriminant Analysis (LDA)

O Linear Discriminant Analysis (LDA) tem como objetivo minimizar as diferenças entre imagens de uma mesma classe e aumentar as diferenças entre imagens de classes diferentes [4]. Para isso, o LDA encontra o ponto de maior separação entre os dados e cria um novo eixo a partir desse ponto, reduzindo a dimensão do espaço.

Podemos compreender melhor a aplicação do LDA em um conjunto de dados por meio de uma representação visual, como mostrado abaixo:

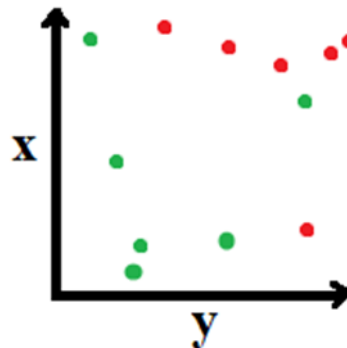


Figura 1: aplicação do LDA em um conjunto de dados por meio de uma representação visual.

Suponhamos um gráfico bidimensional com os eixos X e Y, e desejamos reduzi-lo para uma dimensão, aumentando a separação entre as categorias X e Y. Uma abordagem seria

ignorar as informações fornecidas por um dos eixos e projetar todos os dados no outro eixo, resultando em algo como:



Figura 2: gráfico bidimensional com os eixos X e Y, reduzido a uma dimensão.

No entanto, essa abordagem não separa as informações tão bem quanto gostaríamos. Porém, quando aplicamos o LDA, obtemos resultados significativamente melhores:

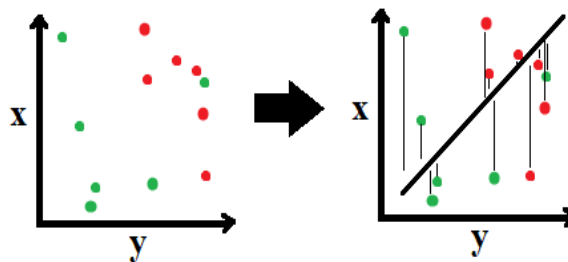


Figura 3: aplicação alternativa do LDA .

Na Figura 2, projetamos apenas as informações do eixo Y no eixo X, o que proporciona uma separação razoável, mas não maximiza a separação entre as duas categorias. Já na Figura 3, criamos uma reta, que se torna o novo eixo, e é nela que realizamos a projeção dos dados.

Para determinar esse novo eixo, utilizamos dois critérios igualmente avaliados: a maior distância entre as médias das classes e a minimização da variação dentro de cada categoria, conhecida como scatter. Podemos descrever esses critérios com a seguinte fórmula:

$$\frac{(\mu_a - \mu_b)^2}{s_a^2 + s_b^2}$$

- Onde μ representa a distância entre as médias das classes e s o scatter.

Em um cenário ideal, buscamos um numerador grande e um denominador pequeno, ou seja, uma grande distância e uma pequena variação. Ambos os critérios são necessários,

pois considerar apenas um deles não proporcionaria uma separação tão eficaz quanto desejamos [4].

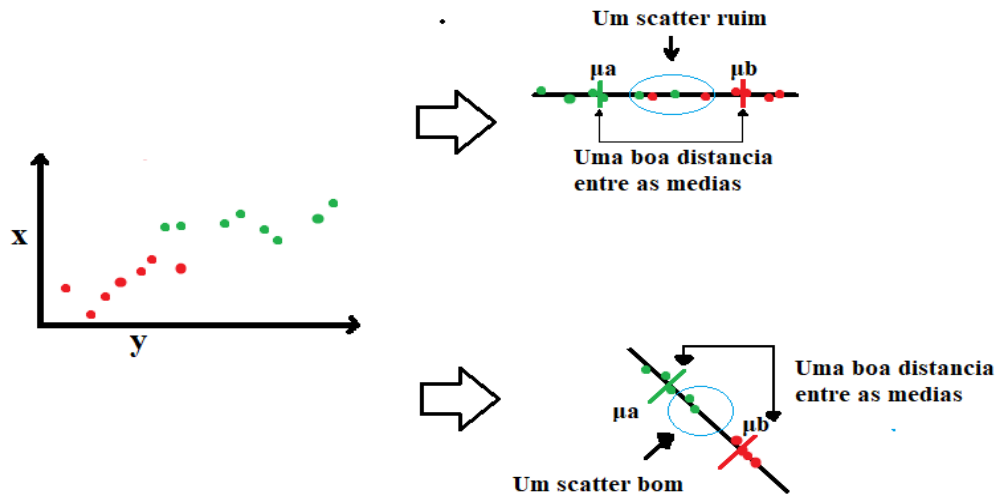


Figura 4: Variação do Scatter.

K Nearest Neighbors (KNN)

O algoritmo K Nearest Neighbors (KNN) busca categorizar um dado aproximando-o de seus "vizinhos". Ele mede a proximidade entre o dado que deseja categorizar e os dados que já são conhecidos, colocando-o na mesma categoria do seu "vizinho mais próximo" [5].

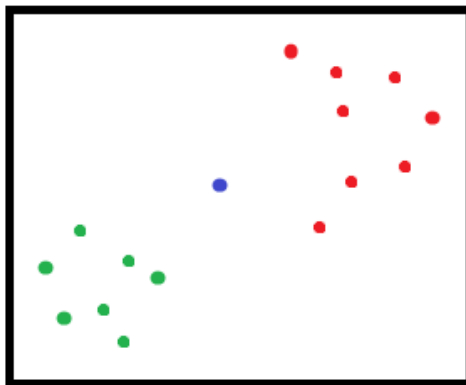


Figura 5: ponto em azul que será categorizado pelo KNN.

Na Figura 5, podemos observar um novo ponto em azul que será categorizado pelo KNN. Para isso, o algoritmo identificará qual outro ponto, ou conjunto de pontos, está mais próximo desse ponto azul no plano e o atribuirá à mesma categoria do ponto mais próximo. O KNN utiliza dois métodos para determinar o vizinho mais próximo:

- **Regressão:** Calcula a distância entre o novo ponto e todos os outros pontos do plano, obtém a média dessas distâncias e seleciona a menor média como a categoria do novo ponto [10].
- **Classificação:** Calcula a distância entre o novo ponto e todos os pontos do plano, identifica a categoria que possui a maior quantidade de pontos próximos ao novo ponto e o classifica nessa categoria [10].

Na Figura 6, na abordagem de regressão, o "vizinho" mais próximo do ponto azul pertenceria à categoria verde, resultando na atribuição da mesma categoria ao ponto azul pelo KNN. Já na abordagem de classificação, o KNN encontraria três "vizinhos" próximos ao ponto azul e o classificaria na mesma categoria dos pontos vermelhos, pois eles possuem mais "vizinhos" próximos em comparação aos pontos verdes, como ilustrado na Figura 7.

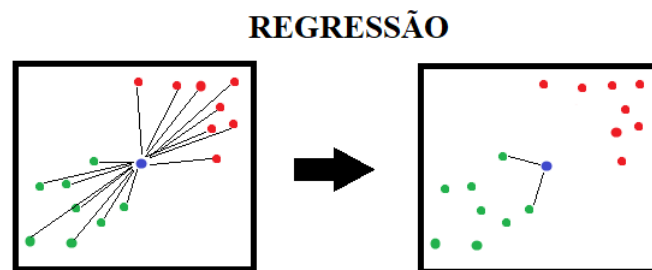


Figura 6: abordagem de regressão do KNN.

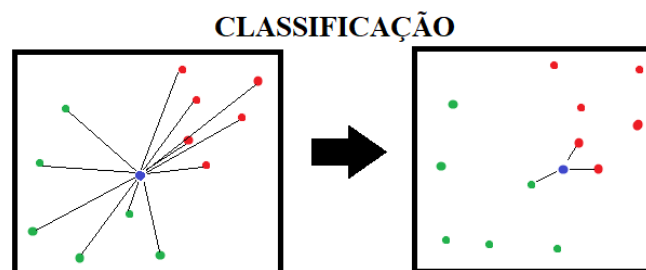


Figura 7: abordagem de classificação do KNN.

Decision Tree Classifier (DTC)

O Decision Tree Classifier (DTC) utiliza a estrutura de árvore para classificar dados de forma simples. O objetivo do DTC é maximizar a separação das informações durante a classificação, dividindo o conjunto de dados em subconjuntos de forma a obter a maior divisão possível [8].

A classificação é realizada com base nas informações presentes nos dados. O DTC constrói uma árvore, começando pelo nó-raiz que representa o conjunto completo de dados. Em cada divisão da árvore, são criados nós-filhos que representam os subconjuntos resultantes. A escolha do atributo para cada nó, seja ele raiz ou filho, é feita usando o Critério de Gini [11].

O Critério de Gini é uma medida que quantifica a impureza dos dados em relação às classes existentes. É calculado pela fórmula $Gini = 1 - (\text{probabilidade de verdadeiro})^2 - (\text{probabilidade de falso})^2$. O DTC utiliza o Critério de Gini para selecionar o atributo que será o nó-raiz e, subsequentemente, para cada nó-filho. Ele realiza uma média ponderada do Critério de Gini dos nós-filhos resultantes de cada atributo e compara os resultados para escolher o atributo que possua o resultado mais próximo de zero [11].

3. MÉTODO

A escolha adequada da base de imagens desempenha um papel fundamental na pesquisa. Neste estudo, utilizamos a reconhecida base de dados Yale Faces [12], composta por 165 imagens distintas, abrangendo 11 categorias e retratando 15 indivíduos diferentes. Optamos por essa base devido à diversidade de categorias disponíveis, que são representativas para todos os indivíduos analisados.

Dentre as várias categorias disponíveis, focamos nossa análise em quatro delas: "center light" (luz centrada), "glasses" (óculos), "happy" (feliz) e "leftlight" (luz à esquerda). Essas categorias foram selecionadas devido às suas características relevantes, que abrangem variações comuns encontradas no reconhecimento facial e são aplicáveis a diversas situações práticas.

Além da escolha cuidadosa da base de imagens, o tratamento adequado das imagens é essencial para garantir uma comparação justa dos resultados. Para isso, adotamos uma série de passos de padronização das imagens. Primeiramente, convertemos as imagens para o padrão de cores RGB (Red Green Blue) a fim de preservar todos os dados de cor relevantes. Em seguida, realizamos a extração dos rostos presentes nas imagens, o que além de economizar recursos computacionais, agiliza os processamentos subsequentes. Por fim,

redimensionamos as imagens dos rostos para um tamanho adequado e uniforme.

Após a conclusão dessas etapas de tratamento de imagens em nossa base de dados, estamos prontos para aplicar os algoritmos de reconhecimento facial e obter resultados significativos.

Base de Imagens

Para a pesquisa, foi selecionada a base de imagens Yale Faces [12], que consiste em um total de 165 imagens diferentes, divididas em 11 categorias e representando 15 indivíduos distintos. A escolha dessa base de dados deve-se à variedade de categorias disponíveis, as quais estão presentes para todos os indivíduos analisados.

Dentre as várias categorias disponíveis, foram selecionadas quatro para análise: "center light" (luz centrada), "glasses" (óculos), "happy" (feliz) e "leftlight" (luz à esquerda). Essas categorias foram escolhidas devido às suas características representativas, que abrangem variações comuns encontradas no reconhecimento facial.

Tratamento de Imagens

O tratamento adequado das imagens desempenha um papel crucial para garantir a comparação justa e precisa dos resultados. A padronização das imagens permite que os algoritmos de reconhecimento facial tenham entradas de dados consistentes, independentemente da classe ou categoria a ser comparada. Para alcançar essa padronização, seguimos os seguintes passos:

Padrão de Cor: Inicialmente, convertemos todas as imagens para o padrão de cores RGB (Red Green Blue). Essa conversão é fundamental para evitar perdas de informações importantes presentes nos valores de cor das imagens.

Extração de Rosto: Em seguida, realizamos a extração precisa dos rostos presentes em cada imagem. Essa etapa não apenas economiza recursos computacionais, mas também acelera os processamentos futuros, pois concentramos nossa análise nas áreas de interesse.

Redimensionamento das Imagens: Após a extração dos rostos, redimensionamos as imagens para um tamanho menor e mais apropriado. Esse redimensionamento uniformiza as imagens, facilitando a aplicação dos algoritmos e garantindo uma análise consistente.

Com a conclusão desses passos de tratamento de imagens em toda a nossa base de dados, estamos prontos para aplicar os algoritmos de reconhecimento facial e obter resultados significativos. Essas etapas nos permitem realizar uma avaliação imparcial e comparativa dos algoritmos, garantindo que as variações nas imagens não afetem indevidamente os resultados finais.

4. RESULTADOS E DISCUSSÃO

A implementação dos conceitos previamente apresentados e descritos é realizada de maneira cuidadosa para garantir a aplicação correta dos algoritmos e a obtenção de resultados confiáveis. Para possibilitar uma análise e comparação adequada entre os algoritmos, estabelecemos uma metodologia padronizada que visa evitar a presença de outliers e facilitar o processo de testes.

O processo de realização dos testes segue determinados parâmetros que foram cuidadosamente estabelecidos:

1. **Hardware Padronizado:** Todos os testes são executados no mesmo hardware, garantindo que as condições de processamento sejam uniformes e que as diferenças nos resultados não sejam influenciadas pelas especificações do dispositivo.
2. **Base de Treinamento e Teste:** Utilizamos a mesma base de dados para treinamento e teste em todos os casos. Essa abordagem assegura que os algoritmos sejam avaliados de forma consistente, sem viés devido a variações nas amostras de dados.
3. **Equivalência de Tamanho:** É importante destacar que as bases de teste e treinamento têm o mesmo tamanho em cada teste realizado. Isso garante que as métricas de desempenho, como a acurácia, sejam comparáveis entre os algoritmos.

Após a garantia desses parâmetros, os testes foram conduzidos e os resultados obtidos foram compilados, como apresentado abaixo:

Tabela 1: resultados dos testes realizados.

algoritmos	Acurácia			
	Centerlight	glasses	happy	leftlight
DTC	0,8	0,73	0,93	0,73
KNN	1	1	1	0,98
LDA	1	1	1	0,93

5. CONSIDERAÇÕES FINAIS (OU CONCLUSÕES)

Ao longo deste artigo, exploramos diversos algoritmos para o reconhecimento facial, compreendendo seu funcionamento e as etapas necessárias para sua implementação. Essa investigação foi realizada com o objetivo de obter uma visão objetiva do desempenho dos diferentes algoritmos em várias situações-problema.

Com base nos resultados obtidos, podemos observar que o KNN apresentou uma acurácia ligeiramente superior ao LDA na categoria "leftlight", enquanto o LDA demonstrou a menor acurácia em todas as categorias analisadas. Podemos afirmar, portanto, que tanto o LDA quanto o KNN podem ser aplicados com sucesso em imagens padronizadas, com iluminação central, pessoas usando óculos e expressões faciais felizes, sem perda significativa de acurácia. No entanto, em casos em que a iluminação está mais à esquerda, o KNN apresenta uma acurácia um pouco melhor em comparação ao LDA. Essa análise nos permite fazer escolhas adequadas dos algoritmos de reconhecimento facial, considerando as características específicas de cada situação.

REFERÊNCIAS

- [1]https://engineering.purdue.edu/~ece624/papers/challapa_facerecognition.pdf (Human and Machine Recognition of Faces: A Survey)
- [2]<https://www.researchpublish.com/upload/book/FACE%20RECOGNITION-3267.pdf> (Face Recognition using Linear Discriminant Analysis)
- [3][http://datajobstest.com/data-science-repo/LDA-Primer-\[Balakrishnama-and-Ganapathiraju\].pdf](http://datajobstest.com/data-science-repo/LDA-Primer-[Balakrishnama-and-Ganapathiraju].pdf) (Linear Discriminant Analysis - A Brief Tutorial)
- [4]<https://www.analyticsvidhya.com/blog/2021/08/a-brief-introduction-to-linear-discriminant-analysis/> (A Brief Introduction to Linear Discriminant Analysis)
- [5]https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm (Wikipedia k-nearest neighbors algorithm)
- [6]<https://isl.stanford.edu/~cover/papers/transIT/0021cove.pdf> (Nearest Neighbor Pattern classification)
- [7]<https://pdfs.semanticscholar.org/baf1/37c478ecea80ad363c447597377abb234089.pdf> (Face Identification Based on K-Nearest Neighbor)
- [8]<https://www.taylorfrancis.com/books/mono/10.1201/9781315139470/classification-regression-trees-leo-breiman> (Classification and Regression Trees)
- [9]<https://www.ise.bgu.ac.il/faculty/liorr/hbchap9.pdf> (Decision Tree)
- [10]https://proceedings.neurips.cc/paper_files/paper/1995/file/fe709c654eac84d5239d1a12a4f71877-Paper.pdf (Discriminant Adaptive Nearest Neighbor Classification and Regression)
- [11]<https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d8c8616e1aefc4dcf13da08104ec5689adb7977> (Exploiting the Cost (In)sensitivity of Decision Tree Splitting Criteria)
- [12] Disponível online em: <http://cvc.cs.yale.edu/cvc/projects/yalefaces/yalefaces.html>