



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

**Análise quantitativa da influência de degradações em  
modelos de aprendizagem profunda utilizados no  
Reconhecimento Facial**

Daniel Moraes da Silva

Monografia apresentada como requisito parcial  
para conclusão do Curso de Bacharelado em Engenharia da Computação

Orientador

Prof. Dr. Flávio de Barros Vidal

Brasília  
2023



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

**Análise quantitativa da influência de degradações em  
modelos de aprendizagem profunda utilizados no  
Reconhecimento Facial**

Daniel Moraes da Silva

Monografia apresentada como requisito parcial  
para conclusão do Curso de Bacharelado em Engenharia da Computação

Prof. Dr. Flávio de Barros Vidal (Orientador)  
CIC/UnB

Prof. Dr. Camilo Chang Dorea    Prof. Dr. Marcelo Marotta  
CIC/UnB    CIC/UnB

Prof. Dr. João Luiz Azevedo de Carvalho  
Coordenador do Curso de Bacharelado em Engenharia da Computação

Brasília, 17 de Julho de 2023

# Dedicatória

Dedico este trabalho a todos aqueles que foram fundamentais em minha jornada, tornando-a uma experiência repleta de aprendizado e crescimento.

À minha família, que sempre esteve ao meu lado, apoiando-me incondicionalmente em cada passo que dei. Vocês são meu porto seguro, minha fonte de amor e inspiração. Sem a presença e o apoio de vocês, esta conquista não seria possível. Dedico-lhes todo o meu sucesso e gratidão.

*2112*

# Agradecimentos

Agradeço sinceramente ao Prof. Dr. Flávio de Barros Vidal, meu orientador, por acreditar em mim e me acolher em seu projeto. Sua orientação experiente e apoio constante foram fundamentais neste trabalho.

Gostaria também de expressar minha profunda gratidão à equipe de pesquisadores envolvida no projeto de reconhecimento facial. O trabalho colaborativo, as discussões enriquecedoras e o esforço conjunto foram absolutamente essenciais para alcançarmos os resultados obtidos. Em especial ao meu parceiro de projeto, Leandro Dias Carneiro por estar mais de perto nessa jornada acadêmica e me ajudar do início ao fim nas pesquisas.

A todos os envolvidos, meu mais sincero agradecimento por compartilharem seus conhecimentos e contribuírem de forma significativa para o desenvolvimento desta pesquisa.

# Resumo

O objetivo deste trabalho é realizar uma análise quantitativa de técnicas de degradação de imagens com a finalidade de aplicação em estudos e pesquisas em reconhecimento facial. A relevância deste trabalho decorre da necessidade de bases de dados contendo imagens degradadas, essenciais para o processo de reconhecimento facial e aprendizado de máquina. Cada imagem degradada será considerada como um objeto de análise para avaliação da qualidade da imagem em um sistema. Foram explorados oito métodos de degradação, e uma análise de dados foi conduzida para cada modelo de reconhecimento facial. Essa análise considerou os estimadores de degradação, suas curvas e o impacto das degradações no desempenho desses modelos. Os resultados obtidos serão utilizados como base para a análise da qualidade da imagem e, posteriormente, para o desenvolvimento de um sistema de aprendizado de máquina profunda.

**Palavras-chave:** Reconhecimento facial, degradação em imagens, Aprendizagem profunda

# Abstract

This work aims to conduct a quantitative analysis of image degradation techniques to apply them in studies and research in facial recognition. The relevance of this work stems from the need for databases containing degraded images, which is essential for facial recognition and machine learning. Each degraded image will be considered an analysis object for evaluating the image quality in a system. Eight degradation methods were explored, and data analysis was conducted for each facial recognition model. This analysis considered the degradation estimators, their curves, and the impact of degradations on the performance of these models. The results obtained will be used as a basis for image quality analysis and, later, for developing a deep machine learning system.

**Keywords:** Facial recognition, Image degradation, Deep learning

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contextualização . . . . .	1
1.2	Motivação . . . . .	2
1.3	Justificativa . . . . .	3
1.4	Objetivos . . . . .	3
1.4.1	Objetivo Geral . . . . .	3
1.4.2	Objetivos Específicos . . . . .	3
1.5	Organização do Trabalho . . . . .	3
<b>2</b>	<b>Fundamentação Teórica</b>	<b>5</b>
2.1	Reconhecimento Facial . . . . .	5
2.1.1	Fluxo de um processo de Reconhecimento Facial Padrão . . . . .	6
2.1.2	Detecção da Face . . . . .	7
2.1.3	Alinhamento da Face . . . . .	8
2.1.4	Representação da Face . . . . .	8
2.1.5	Comparação da Face (Cálculo de Similaridade) . . . . .	9
2.2	Tarefas do Reconhecimento Facial . . . . .	11
2.2.1	Identificação da Face . . . . .	12
2.2.2	Verificação da Face . . . . .	12
2.3	Aprendizagem Profunda . . . . .	12
2.4	Principais Modelos de Reconhecimento Facial Baseado em Aprendizado Profundo . . . . .	13
2.5	Degradações . . . . .	15
2.5.1	Borramento Gaussiano . . . . .	15
2.5.2	Ruído Gaussiano . . . . .	16
2.5.3	Redimensionado . . . . .	17
2.5.4	Escurecimento . . . . .	17
2.5.5	Desfoque de Movimento . . . . .	18
2.5.6	Brilho . . . . .	18

2.5.7	Ruído de Sal e Pimenta . . . . .	18
2.5.8	Compressão JPEG . . . . .	19
<b>3</b>	<b>Metodologia Proposta</b>	<b>20</b>
3.1	Levantamento Bibliográfico das Degradações . . . . .	20
3.2	Modelos de Reconhecimento Facial . . . . .	21
3.3	Bases de Faces . . . . .	22
3.4	Quantificar Degradações . . . . .	23
3.5	Efeitos das Degradações . . . . .	24
<b>4</b>	<b>Resultados</b>	<b>25</b>
4.1	Base de imagens - <i>Labeled Faces in the Wild</i> . . . . .	25
4.2	Descrição dos Testes . . . . .	25
4.3	Descrição dos Gráficos Elaborados . . . . .	27
4.3.1	Precision P+ . . . . .	27
4.3.2	Precision P- . . . . .	27
4.3.3	Recall P+ . . . . .	27
4.3.4	Recall P- . . . . .	27
4.4	Teste 1 - Imagem degradada x degradada . . . . .	28
4.5	Teste 2 - Imagem original x degradada . . . . .	31
4.6	Teste 3 - Imagem original x mesma imagem original até intensidade máxima	33
4.6.1	Dos resultados . . . . .	34
<b>5</b>	<b>Conclusões</b>	<b>36</b>
	<b>Referências</b>	<b>38</b>
	<b>Anexo</b>	<b>41</b>
<b>I</b>	<b>Resultados Completos</b>	<b>42</b>
I.1	No contexto do Teste 1	
Imagem degradada x degradada . . . . .		42
I.1.1	Borramento Gaussiano . . . . .	43
I.1.2	Ruído Gaussiano . . . . .	47
I.1.3	Redimensionado . . . . .	48
I.1.4	Escurecimento . . . . .	49
I.1.5	Desfoque de Movimento . . . . .	51
I.1.6	Ruído de Sal e Pimenta . . . . .	53
I.1.7	Compressão JPEG . . . . .	55



I.2	No contexto do Teste 2	
	Imagem original x degradada . . . . .	57
I.2.1	Ruído Gaussiano . . . . .	57
I.2.2	Borramento Gaussiano . . . . .	59
I.2.3	Escurecimento . . . . .	61
I.2.4	Desfoque de Movimento . . . . .	63
I.2.5	Ruído de Sal e Pimenta . . . . .	65
I.2.6	Compressão JPEG . . . . .	67
I.3	No contexto do Teste 3	
	Imagem original x mesma imagem original . . . . .	69
I.3.1	Ruído Gaussiano . . . . .	69
I.3.2	Borramento Gaussiano . . . . .	71
I.3.3	Escurecimento . . . . .	73
I.3.4	Desfoque de Movimento . . . . .	75
I.3.5	Ruído de Sal e Pimenta . . . . .	77
I.3.6	Compressão JPEG . . . . .	79

# Lista de Figuras

2.1	Fluxogram padrão de um sistema de reconhecimento facial . . . . .	7
2.2	Pose facial observados na ISO/IEC 39794-5 . . . . .	8
2.3	Vetor de representação de uma imagem . . . . .	9
2.4	Exemplo de funcionamento de uma função de perda do tipo <i>triplet loss</i> . . .	10
2.5	Borramento Gaussiano . . . . .	15
2.6	Ruído Gaussiano . . . . .	16
2.7	Redimensionado . . . . .	17
2.8	Escurecimento . . . . .	17
2.9	Desfoque de Movimento . . . . .	18
2.10	Brilho . . . . .	18
2.11	Ruído de Sal e Pimenta . . . . .	19
2.12	Compressão JPEG . . . . .	19
3.1	Fluxograma da Metodologia . . . . .	20
4.1	Ilustração de combinações para o Teste 1. . . . .	26
4.2	Melhor desempenho - Facenet512 - Compressão JPEG . . . . .	29
4.3	Pior desempenho - ArcFace Borramento Gaussiano . . . . .	30
4.4	Melhor desempenho - Facenet512 Ruído Gaussiano . . . . .	32
4.5	Pior desempenho - VGG-Face Ruído Sal e Pimenta . . . . .	32
4.6	Melhor desempenho - VGG-Face Compressão JPEG . . . . .	34
4.7	Pior desempenho - OpenFace Escurecimento . . . . .	34
I.1	Borramento Gaussiano - ArcFace . . . . .	43
I.2	Borramento Gaussiano - DeepFace . . . . .	45
I.3	Borramento Gaussiano - Dlib . . . . .	46
I.4	Borramento Gaussiano - Facenet512 . . . . .	47
I.5	Ruído Gaussiano - ArcFace . . . . .	48
I.6	Redimensionado - ArcFace . . . . .	49
I.7	Escurecimento - DeepFace . . . . .	50

I.8	Escurecimento - VGG-Face . . . . .	51
I.9	Desfoque de Movimento - ArcFace . . . . .	52
I.10	Desfoque de Movimento - VGG-Face . . . . .	53
I.11	Ruído de Sal e Pimenta - ArcFace . . . . .	54
I.12	Ruído de Sal e Pimenta - VGG-Face . . . . .	55
I.13	Compressão JPEG - VGG-Face . . . . .	56
I.14	Ruído Gaussiano - ArcFace . . . . .	57
I.15	Ruído Gaussiano - VGG-Face . . . . .	58
I.16	Borramento Gaussiano - DeepFace . . . . .	59
I.17	Borramento Gaussiano - VGG-Face . . . . .	60
I.18	Escurecimento - DeepFace . . . . .	61
I.19	Escurecimento - VGG-Face . . . . .	62
I.20	Desfoque de Movimento - DeepFace . . . . .	63
I.21	Desfoque de Movimento - VGG-Face . . . . .	64
I.22	Ruído de Sal e Pimenta - DeepFace . . . . .	65
I.23	Ruído de Sal e Pimenta - VGG-Face . . . . .	66
I.24	Compressão JPEG - DeepFace . . . . .	67
I.25	Compressão JPEG - VGG-Face . . . . .	68
I.26	Ruído Gaussiano - ArcFace . . . . .	69
I.27	Ruído Gaussiano - VGG-Face . . . . .	70
I.28	Borramento Gaussiano - Facenet512 . . . . .	71
I.29	Borramento Gaussiano - OpenFace . . . . .	72
I.30	Escurecimento - DeepID . . . . .	73
I.31	Escurecimento - OpenFace . . . . .	74
I.32	Desfoque de Movimento - Facenet512 . . . . .	75
I.33	Desfoque de Movimento - OpenFace . . . . .	76
I.34	Ruído de Sal e Pimenta - Facenet512 . . . . .	77
I.35	Ruído de Sal e Pimenta - OpenFace . . . . .	78
I.36	Compressão JPEG - OpenFace . . . . .	79
I.37	Compressão JPEG - SFace . . . . .	80

# Lista de Tabelas

4.1	Tabela de testes 1 . . . . .	28
4.2	Tabela de testes 2 . . . . .	31
4.3	Tabela de testes 3 . . . . .	33

# Capítulo 1

## Introdução

### 1.1 Contextualização

O reconhecimento facial em imagens desempenha um papel crucial em diversas áreas, tendo uma importância crescente na sociedade atual. Essa tecnologia permite a identificação e autenticação de indivíduos com base em suas características faciais únicas, oferecendo uma forma conveniente e segura de acesso a dispositivos, sistemas e serviços[1]. Além disso, o reconhecimento facial tem aplicações abrangentes em segurança, como monitoramento de locais públicos, detecção de fraudes e prevenção de crimes. Também é utilizado em setores como saúde, com diagnóstico de doenças dermatológicas, e em entretenimento, como em jogos e aplicativos de realidade aumentada.

A aplicação de técnicas de degradação no campo de processamento de imagens desempenha um papel fundamental na avaliação e aprimoramento dos sistemas de reconhecimento facial [2]. Essas técnicas envolvem a simulação controlada de distorções e imperfeições que podem ocorrer em imagens do mundo real, como borramento, ruído, compressão e variações de iluminação. Essas simulações têm o propósito de investigar como os modelos de aprendizagem profunda reagem a essas condições desafiadoras e como a acurácia do reconhecimento facial é afetada. Ao aplicar essas técnicas de degradação, pode-se obter resultados valiosos sobre a robustez e a capacidade de generalização dos modelos, bem como desenvolver estratégias para melhorar o desempenho desses sistemas em cenários realistas e variados. O presente projeto visa justamente realizar uma análise quantitativa dessas técnicas de degradação, contribuindo assim para o avanço no entendimento e aprimoramento dos sistemas de reconhecimento facial baseados em aprendizado profundo.[3]

O uso de modelos de aprendizagem profunda no reconhecimento facial tem revolucionado essa área, proporcionando avanços significativos em termos de precisão e desempenho. Esses modelos são capazes de aprender automaticamente a partir de grandes volumes

de dados de treinamento, identificando padrões complexos e representando características faciais de forma eficiente. Ao contrário dos métodos tradicionais, que dependem de algoritmos manuais e extração de características específicas, os modelos de aprendizagem profunda podem aprender características de alto nível diretamente dos dados, permitindo uma representação mais abrangente e robusta das faces[4].

Essa abordagem possibilita uma melhor generalização e adaptação a diferentes condições de iluminação, expressões faciais e variações de pose. Além disso, os modelos de aprendizagem profunda têm a capacidade de aprender representações hierárquicas, capturando informações em diferentes níveis de abstração e melhorando a capacidade de discriminação e reconhecimento de indivíduos. Essa combinação de capacidades torna os modelos de aprendizagem profunda uma escolha poderosa e promissora para aplicações de reconhecimento facial em diversos domínios[5].

Este projeto tem como objetivo realizar uma análise quantitativa dos impactos causados pelas degradações em sistemas de reconhecimento facial, especialmente no âmbito da pesquisa em reconhecimento facial baseado em aprendizado profundo. A relevância dessa investigação é respaldada pela necessidade de se obter conjuntos de imagens degradadas, fundamentais para aprimorar o processo de reconhecimento facial e o aprendizado de máquina. Nessa abordagem, cada imagem degradada se torna um alvo de análise, permitindo uma avaliação detalhada da qualidade da imagem no contexto desse sistema de reconhecimento facial.

Foram abordados oito modelos de reconhecimento facial apresentados em [6]. Realizou-se uma análise de dados para cada método, considerando os estimadores de degradação, suas curvas e a influência das degradações no reconhecimento facial. Esses resultados servirão como base para a análise da qualidade de imagem e, posteriormente, para o desenvolvimento de um sistema de aprendizado profundo de máquina.

## 1.2 Motivação

A motivação para este trabalho é a necessidade de analisar quantitativamente as técnicas de degradação de imagens para aplicação em estudos e pesquisas em reconhecimento facial, com foco na área pericial e na gestão da criminalística. A contribuição deste trabalho está relacionada à análise dos efeitos das degradações em sistemas de reconhecimento facial baseado em aprendizado profundo, bem como a geração de bases de imagens degradadas, essenciais para avaliar a acurácia e a qualidade das imagens nesse sistema de reconhecimento facial.

## 1.3 Justificativa

A justificativa para este projeto reside na importância de compreender o efeito das degradações nas técnicas de reconhecimento facial, especialmente em ambientes forenses. A análise quantitativa das degradações é fundamental para identificar os principais fatores que influenciam a acurácia dos modelos de aprendizado profundo. Compreender o quanto essas degradações afetam o desempenho dos sistemas é crucial para o desenvolvimento de soluções mais robustas e confiáveis.

## 1.4 Objetivos

### 1.4.1 Objetivo Geral

O objetivo geral deste trabalho é analisar e quantificar o impacto das degradações nas técnicas de reconhecimento facial que utilizam modelos de aprendizado profundo. O foco está na identificação das principais degradações e na avaliação de seu efeito na acurácia dos sistemas.

### 1.4.2 Objetivos Específicos

Para alcançar o objetivo geral, são definidos os seguintes objetivos específicos:

- Identificar as principais degradações que afetam diretamente a acurácia dos sistemas de reconhecimento facial baseados em modelos de aprendizado profundo.
- Quantificar o impacto das degradações na medida de desempenho dos sistemas de reconhecimento facial.
- Desenvolver uma metodologia para análise e quantificação das degradações em imagens faciais.
- Propor melhorias e recomendações para mitigar os efeitos negativos das degradações nos sistemas de reconhecimento facial.

## 1.5 Organização do Trabalho

Este trabalho está organizado em capítulos da seguinte forma:

- **Capítulo 1 - Introdução:** neste capítulo, são apresentados o contexto, a motivação, a justificativa, os objetivos e a organização do trabalho.

- **Capítulo 2 - Fundamentação Teórica:** aborda os conceitos e as técnicas fundamentais relacionados ao reconhecimento facial, aprendizado profundo e degradação de imagens.
- **Capítulo 3 - Metodologia:** são apresentados os detalhes da metodologia adotada para realizar a análise quantitativa das degradações em modelos de reconhecimento facial.
- **Capítulo 4 - Resultados:** os resultados obtidos na análise das degradações são apresentados e discutidos neste capítulo.
- **Capítulo 5 - Conclusão:** neste capítulo, são apresentadas as conclusões do trabalho, incluindo as contribuições, as limitações e as sugestões para trabalhos futuros.



# Capítulo 2

## Fundamentação Teórica

Neste capítulo será apresentado uma breve fundamentação que compõe todo o arcabouço teórico sobre um sistema de reconhecimento facial, aprendizagem profunda e seus modelos utilizados, degradações, bem como os conceitos relativo ao âmbito de processamento de imagens.

### 2.1 Reconhecimento Facial

No campo da segurança biométrica existem vários sistemas biométricos para identificar ou confirmar a identidade de uma pessoa, como a impressão digital, o reconhecimento por voz, retina ocular ou por iris, assinatura, palma da mão, entre outros. O sistema reconhecimento facial é uma forma de identificar uma pessoa usando seu rosto e atualmente é um dos sistemas mais seguros e utilizados em diversas aplicações.

Os sistemas biométricos baseados em Reconhecimento Facial em geral exploram características relacionadas à localização e forma de atributos faciais como olhos, sobrancelhas, nariz, lábios entre outros. Os desafios surgem quando o sistema é submetido a condições adversas, como variação de luminosidade, degradações nas imagens e baixa qualidade das imagens captadas, entre outras. Devido a isso, esse tipo de sistema biométrico ainda enfrenta imperfeições que o distancia de uma forma definida, porém é um campo amplamente aberto para melhorias [7].

O Reconhecimento de faces vem ganhando bastante notoriedade nos últimos anos, principalmente com o advento de câmeras digitais presentes em boa parte do cotidiano das pessoas, como em seus smartphones, áreas de segurança ou até mesmo em locais públicos. Sendo assim o Reconhecimento Facial por imagens, sejam elas por fotografias ou vídeos, é amplamente utilizado.

Na prática pericial é uma das aplicações que utiliza reconhecimento por fotografia e que ganhou bastante destaque devido a alta coleta de matérias oriundas de câmeras de

vigilância e bancos de dados com imagens faciais de pessoas para comparação e verificação de suas identidades. Além do amplo uso na identificação de pessoas, ele vem sendo usado em práticas forenses para a reconstrução de rostos a partir de suas formas faciais detectadas e mapeadas [8].

Hoje em dia, boa parte das técnicas fazem uso de tecnologias baseadas em inteligência artificial(IA)[9]. Essas técnicas têm revolucionado a forma como interagimos com dispositivos e sistemas de segurança, oferecendo uma maneira eficiente e precisa de identificar e autenticar indivíduos por meio de características faciais únicas. A IA possibilita que os algoritmos aprendam a reconhecer padrões e características específicas nos rostos das pessoas, permitindo a criação de modelos computacionais capazes de detectar e reconhecer rostos em tempo real com uma taxa de sucesso cada vez maior. Além disso, ela também permite aprimorar a precisão do reconhecimento facial, reduzindo erros e aumentando a confiabilidade dos sistemas. No entanto, é importante ponderar sobre questões éticas e de privacidade relacionadas ao uso indiscriminado dessa tecnologia, garantindo que seu desenvolvimento e aplicação sejam pautados por diretrizes éticas e regulamentações adequadas.

### **2.1.1 Fluxo de um processo de Reconhecimento Facial Padrão**

Uma das tecnologias mais utilizadas para o processo de reconhecimento facial é o Aprendizado Profundo (*Deep Learning*). O Aprendizado profundo é uma subárea da inteligência Artificial que se concentra em treinar redes neurais artificiais profundas inspiradas no funcionamento do cérebro humano, para aprender a partir de dados e realizar tarefas complexas. Essas redes consistem em várias camadas de unidades interconectadas que transformam gradualmente os dados de entrada em representações mais abstratas. Assim, no contexto deste projeto, o processo de reconhecimento divide-se em quatro etapas de execução sobre a face da pessoa [10] e descrito na Figura 2.1:

1. Detecção
2. Alinhamento
3. Representação
4. Comparação (Cálculo de Similaridade)

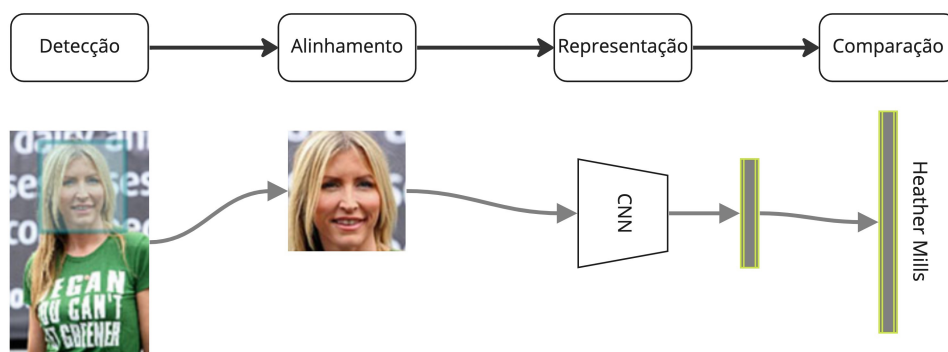


Figura 2.1: Fluxogram padrão de um sistema de reconhecimento facial

### 2.1.2 Detecção da Face

Nesta etapa, o objetivo é a detecção de todas as faces existentes em uma imagem. Após essa varredura o algoritmo retorna as coordenadas dos boxes de demarcações faciais e com seus respectivos graus de confiança.

Por um breve tempo na academia, umas das formas de detecção mais usada era puramente estatística. Esse processo de detecção consistia basicamente de marcações vetoriais as quais listavam vetores de pesos obtidos a partir de componentes Eigenfaces. [11]. Historicamente, esses sistemas eram testados em bases pequenas, porém, apresentavam escores precisos e com solidez suficiente para serem adaptadas ao longo dos avanços computacionais futuros. Atualmente esse grau de precisão evoluiu significativamente, devido os avanços computacionais envolvendo aprendizado profundo e bases (*datasets*) mais robustas e sofisticadas.

No decorrer dos anos, as tecnologias foram evoluindo até que em 2012, uma solução de reconhecimento facial baseada na Rede AlexNet [12] ganha destaque na academia, inovando com a utilização do método de aprendizagem profunda, especificamente em Redes Neurais Convolucionais (CNN - *Convolutional Neural Network*). Com o advento das redes neurais, o nível de abstração e complexidade elevaram, trazendo mais amadurecimento no campo e gerando novos métodos de detecção mais sofisticados e precisos.

Assim, atualmente muitos trabalhos destacam que os algoritmos de detecção de faces baseados em CNNs são robustos e extremamente precisos na detecção quanto a variação de ângulo facial, iluminação e escala. Entretanto, em alguns aspectos como em imagens de baixa resolução e oclusões de face, as CNNs ainda carecem de mais pesquisa e avanços.

### 2.1.3 Alinhamento da Face

O alinhamento da face visa levantar as informações de pose do rosto, tamanho da imagem, bem como suas propriedades fotográficas como a iluminação e escala do cinza. Nessa etapa é capturado vários pontos-chave (*facial landmarks*) que serão utilizados pelo algoritmo nas etapas posteriores.

Vale ressaltar que a pose facial é comumente representada pelos ângulos de inclinação (pitch), guinada (yaw) e rotação (roll), como apresentado na Figura 2.2, e seus critérios de alinhamento são definidos pela ISO/IEC 39794-5.

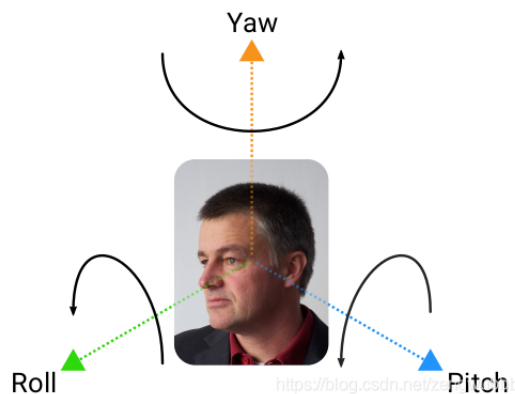


Figura 2.2: Pose facial observados na ISO/IEC 39794-5

### 2.1.4 Representação da Face

Esse é o ponto central de um sistema de reconhecimento facial, executado a partir dos dados extraídos na etapa de alinhamento. A representação, como o nome sugere, consiste em representar por meios de valores numéricos a face da imagem em questão. Desta forma é possível calcular a similaridade entre duas faces onde esse cálculo é comumente feito usando *Aprendizado profundo*

A camada de representação é executada para decidir em qual classe a imagem pertence. Essa classificação é relevante, pois esses valores são passados entre as camadas de um sistema de reconhecimento em que a última camada desse processo decide a saída com base nos valores obtidos da penúltima camada. Assim, a penúltima camada possui os valores que representam as características da imagem. Portanto, conforme ilustrado na Figura 2.3 um vetor de valores se dá pela análise extraída da penúltima camada de uma CNN (para classificação de imagem).

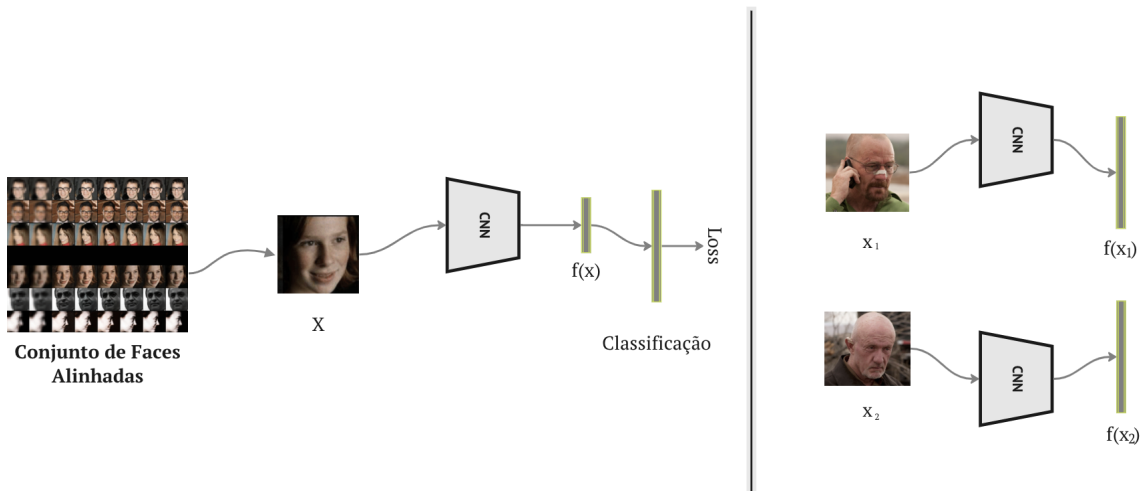


Figura 2.3: Vetor de representação de uma imagem

### 2.1.5 Comparação da Face (Cálculo de Similaridade)

Após a execução das representações numéricas das faces captadas é feito o cálculo de similaridade, onde será determinado se é a mesma pessoa ou não.

O cálculo de similaridade é uma etapa importante no processo de reconhecimento facial, pois permite comparar as características de uma face com as de outras faces armazenadas no banco de dados de referência, determinando assim a pessoa. Em geral, a similaridade é medida usando medições numéricas de distância Euclidiana ou de distância Mahalanobis.

A Distância de Mahalanobis é uma medida de distância que leva em conta a covariância entre as características faciais. Isso é extremamente útil quando as características faciais são altamente correlacionadas, o que pode afetar o resultado da comparação[13]. O cálculo é definido como:

$$D = \sqrt{(x_i - v)^T S^{-1}(x_i - v)} \quad (2.1)$$

Em que, dado um conjunto de amostras  $X = x_1, x_2, \dots, x_n \in \mathbb{R}^{d \times n}$ , em que sua média é descrito como  $v \in \mathbb{R}^d$  e  $S^{-1}$  é a inversa da matriz positiva definida de covariância das amostras.

A distância Euclidiana foi bastante utilizada por um longo tempo, seu cálculo pode ser definido conforme a Equação 2.2 porém com vetores definidos em um espaço amostral não acumulativo e em módulo da média das matrizes.

$$D = \sqrt{(x_i - v)^T (x_i - v)} = \|x_i - v\| \quad (2.2)$$

## Funções de distância em um espaço Euclidiano

Em uma análise mais apurada, as funções euclidianas comportam as imagens em um espaço Euclidiano em que busca reduzir as variações da mesma pessoa ao passo que aumenta as variações de pessoas diferentes. As principais funções desse grupo são, de acordo com [14]:

- *Contrastive loss*;
- *triplet loss*;

A *Contrastive loss* é empregada para treinamento utilizando pares de imagens positivas idênticas ou pares de imagens negativas distintas. Ela considera as discrepâncias em distâncias absolutas como métrica bem definidas. Por outro lado, a função de *triplet loss* é aplicada no treinamento envolvendo três imagens: uma âncora (*baseline*), comparada com uma imagem positiva idêntica (minimização) e uma imagem negativa diferente (maximização), tendo as distâncias relativas como medida canônica. Assim, no processo de treinamento, as três imagens (âncora, positiva e negativa) são tratadas como uma entidade única, sendo concatenadas para essa finalidade.[15].

Conforme ilustrado na Figura 2.4 é possível notar que a distância entre a âncora e o exemplo positivo deve ser menor do que entre a âncora e o negativo.

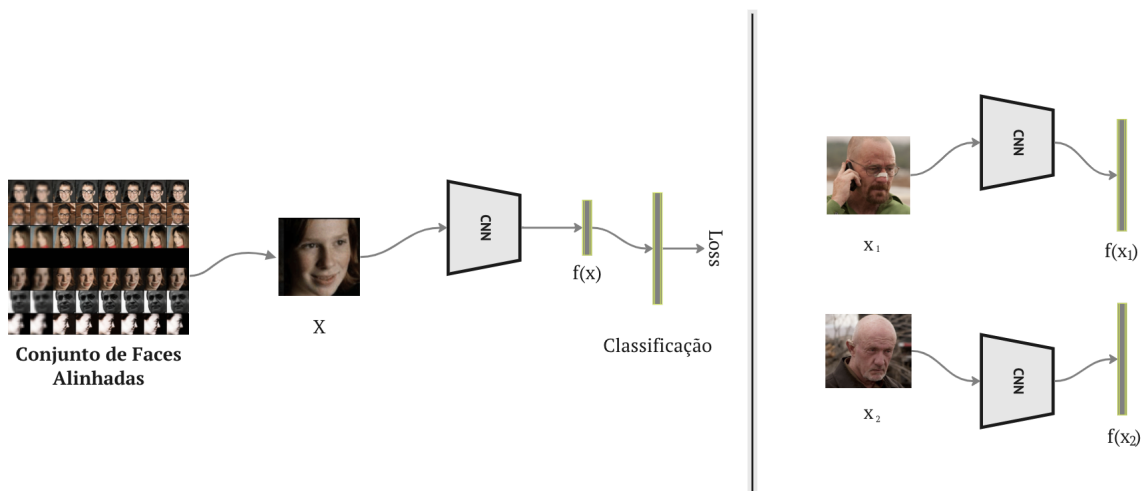


Figura 2.4: Exemplo de funcionamento de uma função de perda do tipo *triplet loss*.

A função de Distância do Cosseno é outra medida comumente utilizada para o cálculo de similaridade em reconhecimento facial. Essa medida leva em consideração o ângulo entre os vetores de características faciais e é especialmente útil quando as variações de

iluminação e escala podem afetar as medições de distância Euclidiana ou de Mahalanobis [16]. A distância do cosseno é definida pela fórmula:

$$\text{Similaridade} = \cos(\theta) = \frac{x_i \cdot v}{|x_i||v|} \quad (2.3)$$

Onde  $x_i$  representa o vetor de características da face a ser comparada e  $v$  é o vetor de características da face de referência no banco de dados. Quanto mais próximo o valor de similaridade estiver de 1, maior será a similaridade entre as faces.

Além dessas medidas de distância, também existem outras funções de similaridade utilizadas em reconhecimento facial. Por exemplo, a função de similaridade baseada na correlação de Pearson mede a similaridade entre as características faciais considerando a correlação linear entre elas. Essa medida é útil quando as características faciais têm uma relação linear significativa.

Cada função de distância ou similaridade tem suas características e aplicabilidades específicas, e a escolha adequada depende das necessidades e dos requisitos do sistema de reconhecimento facial em questão.

## 2.2 Tarefas do Reconhecimento Facial

Existem várias tarefas relacionadas ao reconhecimento facial, algumas, embora já vista nas sessões anteriores, compondo parte do processo de reconhecimento facial padrão [17]. Outrossim, as principais tarefas em reconhecimento são:

- Detecção facial: Encontrar a localização dos rostos em uma imagem ou vídeo;
- Alinhamento facial: Rotacionar e escalar o rosto;
- Extração de características faciais: Extrair características específicas do rosto;
- Verificação Facial: Determinar se as faces contidas nas imagens correspondem a mesma pessoa;
- Identificação Facial: Dado um conjunto de referência de rostos previamente identificados, identificar a qual indivíduo corresponde um rosto desconhecido;
- Análise de expressões faciais: Avaliar as expressões faciais em um rosto, como sorriso, raiva, surpresa, etc;
- Reconstrução Facial 3D: Reconstruir uma representação tridimensional de um rosto a partir de imagens bidimensionais;
- Edição de rostos: Alterar características faciais de um rosto, como cor dos olhos, formato da mandíbula ou nariz, etc;

Para o escopo deste trabalho, a problemática de reconhecimento facial pode ser dividido em duas tarefas, conforme estudos relacionados a eficácia do reconhecimento facial [18], sendo a identificação da face e a verificação da face. Ambas realizam o procedimento de cálculo da representação da face e a comparação com outra imagem. Até então, é notório que em muitos casos, essas são tarefas similares, pois realizam o mesmo pipeline exposto anteriormente: detecção, alinhamento, representação e comparação. Entretanto, o objetivo final das tarefas difere, conforme detalhado a seguir.

### 2.2.1 Identificação da Face

Como apresentado na Sessão 2.2, o conceito mais aprofundado de identificação da face consiste em dado uma base de imagens contendo faces e uma imagem em questão contendo apenas uma face, ambas submetidas a um sistema de reconhecimento facial, o qual deve indicar a imagem pertencente à base que mais se assemelha com a imagem questionada. Assim, para esse procedimento, é realizado o cálculo de similaridade entre a representação da imagem questionada e todas as representações das imagens pertencentes à base, e finalmente, selecionar a imagem com o maior grau de similaridade (ou menor distância facial). Este procedimento é conhecido como 1-N ("um-para-muitos") ou *face matching*.

### 2.2.2 Verificação da Face

Na verificação o principal objetivo é determinar se em duas imagens distintas, correspondem a uma mesma pessoa. Esse procedimento é feito comparando as representações entre as duas imagens e verificando o grau de similaridade a partir de um limiar (*threshold*). Assim, se o valor da distância facial estiver abaixo do limiar, as imagens são consideradas da mesma pessoa. Caso contrário, são pessoas diferentes. Este procedimento é denominado de 1-1 ("um-para-um").

## 2.3 Aprendizagem Profunda

No aprendizagem Profunda suas redes neurais são capazes aprender representações complexas e abstratas de dados, o que as torna muito úteis para resolver vários tipos de problemas, incluindo classificação, detecção de objetos e reconhecimento facial.

O aprendizado profundo é fundamentado em uma variedade de algoritmos e técnicas, abrangendo desde as redes neurais de propagação direta [19], até as redes neurais convolucionais (CNNs) e as redes neurais recorrentes (RNNs). Esses algoritmos operam através da iteração de treinamentos da rede neural utilizando conjuntos extensos de dados, ajustando os pesos das conexões entre suas unidades. Esse processo visa minimizar a perda ou



erro nas previsões, aprimorando assim a capacidade da rede de aprender representações significativas dos dados.

O aprendizado profundo tem sido amplamente aplicado com sucesso em vários domínios e parte de sua popularidade se dá ao fato da disponibilidade de bibliotecas e *frameworks* de softwares, como TensorFlow, tornou o campo de estudo acessível para muitos desenvolvedores e pesquisadores.

No entanto, é importante destacar que essa subárea ainda apresenta algumas limitações, como a necessidade de grandes quantidades de dados para treinamento e a dificuldade de interpretar as decisões tomadas pelas redes neurais. Além disso, a falta de diversidade e equidade nos dados utilizados para treinar os modelos pode levar a discriminação e problemas éticos.

## 2.4 Principais Modelos de Reconhecimento Facial Baseado em Aprendizado Profundo

Nesta seção, serão apresentados os principais modelos utilizados neste trabalho, que foram selecionados por representarem o estado da arte em suas respectivas épocas. Cada modelo é acompanhado de uma breve descrição sobre sua estrutura, a base de dados utilizada em seu treinamento e uma referência bibliográfica para consulta.

- **VGG-Face:** O modelo VGG-Face é uma rede neural convolucional desenvolvida pela Visual Geometry Group (VGG) da Universidade de Oxford. Ele foi treinado em uma ampla variedade de imagens faciais para realizar tarefas de reconhecimento facial [20].
- **Google FaceNet:** O FaceNet, desenvolvido pelo Google, é um modelo de aprendizado profundo baseado em redes neurais convolucionais e de siameses. Ele utiliza a técnica de *embeddings* para mapear imagens faciais em um espaço vetorial, onde distâncias entre os vetores representam similaridades entre as faces [21].
- **OpenFace:** O OpenFace é a primeira ferramenta de código aberto capaz de detecção de marcos faciais, estimativa de pose da cabeça, reconhecimento de unidades de ação facial e estimativa de olhar. Os algoritmos de visão computacional que compõem o núcleo do OpenFace demonstram resultados de ponta. Além disso, esse modelo é capaz de desempenho em tempo real e pode ser executado a partir de uma simples webcam, sem a necessidade de hardware especializado. Por fim, o OpenFace permite integração fácil com outras aplicações e dispositivos [22].

- **Facebook DeepFace:** O DeepFace, desenvolvido pelo Facebook, é um modelo de reconhecimento facial que utiliza redes neurais convolucionais para realizar a tarefa de reconhecimento facial em larga escala. Ele foi treinado em uma vasta quantidade de dados e obteve um desempenho comparável ao de especialistas humanos [23].
- **DeepID:** O DeepID é um modelo de reconhecimento facial proposto por [24] que utiliza redes neurais convolucionais para extrair características faciais e realizar a correspondência entre rostos. Ele foi treinado em uma grande base de dados e alcançou resultados significativos em várias tarefas de reconhecimento facial [24].
- **ArcFace:** ArcFace é um algoritmo avançado de reconhecimento facial que utiliza aprendizado de máquina. Ele foi desenvolvido com o objetivo de melhorar a precisão e a robustez do reconhecimento facial, lidando com desafios como variações de pose, iluminação e expressões faciais. O algoritmo mapeia as características discriminativas da face em um espaço de alta dimensão, onde faces da mesma pessoa ficam próximas e faces de pessoas diferentes ficam afastadas. Uma função de perda angular, chamada "*ArcMarginLoss*", é usada no treinamento para aumentar a separabilidade entre as classes de faces. O modelo tem alcançado excelentes resultados em conjuntos de dados de referência e benchmarks, superando outros algoritmos populares. Sua precisão e robustez o tornam adequado para aplicações como autenticação biométrica, identificação em massa, vigilância por vídeo e sistemas de segurança [25].
- **Dlib:** A análise do algoritmo de reconhecimento facial Dlib tem sido amplamente abordada na literatura devido à sua eficácia e desempenho impressionantes. O Dlib é uma biblioteca de código aberto que oferece diversas funcionalidades para o processamento de imagens, incluindo a detecção e o reconhecimento de faces.  
Uma das principais características do Dlib é sua capacidade de detecção de pontos de referência faciais, também conhecidos como landmarks. Esses landmarks são pontos-chave na face, como olhos, nariz, boca e contornos faciais, que fornecem informações importantes para a identificação e análise facial. A detecção desses pontos é fundamental para várias tarefas, como reconhecimento de expressões faciais e estimativa de pose da cabeça [26].
- **SFace:** O SFace é um modelo de reconhecimento facial que combina redes neurais convolucionais e de siameses com técnicas de atenção espacial. Ele foi desenvolvido para melhorar a precisão e a robustez do reconhecimento facial em condições desafiadoras, como variações de iluminação, expressões faciais e oclusões parciais. Ao utilizar essa abordagem, o SFace é capaz de capturar características faciais relevantes e distintivas, tornando-se eficaz na identificação precisa de indivíduos. Sua

combinação de técnicas o torna adequado para aplicações que exigem reconhecimento facial preciso e robusto em sistemas de segurança e monitoramento de vídeo [27].

Esses modelos representam avanços significativos no campo do reconhecimento facial e foram escolhidos como referência para o presente trabalho.

## 2.5 Degradações

As degradações são manipulações aplicadas a imagens que podem afetar sua qualidade, clareza e informações visuais. Nesta seção, abordaremos as seguintes degradações:

1. Borramento Gaussiano
2. Ruído Gaussiano
3. Redimensionado
4. Escurecimento
5. Desfoque de Movimento
6. Brilho
7. Ruído de Sal e Pimenta
8. Compressão PEG

### 2.5.1 Borramento Gaussiano

O borramento gaussiano, também conhecido como desfoque de média ponderada, é uma técnica comum usada para suavizar imagens, eliminando detalhes finos e reduzindo o ruído. Ele aplica uma média ponderada dos valores dos pixels vizinhos para cada pixel na imagem. Esse tipo de desfoque segue a distribuição gaussiana para determinar a ponderação dos pixels vizinhos. [28]



Figura 2.5: Borramento Gaussiano

## 2.5.2 Ruído Gaussiano

O ruído gaussiano é um tipo de ruído aleatório que segue uma distribuição normal, também conhecida como distribuição gaussiana. É adicionado às imagens com o objetivo de simular imperfeições ou distorções encontradas em várias condições do mundo real. O ruído gaussiano pode afetar a nitidez, a clareza e a qualidade geral da imagem. [29]



Figura 2.6: Ruído Gaussiano

### 2.5.3 Redimensionado

O redimensionamento é um processo de redução do tamanho da imagem, geralmente aplicado para diminuir a quantidade de dados necessários para representar uma imagem. Ao reduzir a resolução, informações finas e detalhes podem ser perdidos, resultando em uma imagem com menos clareza e qualidade visual. [30]

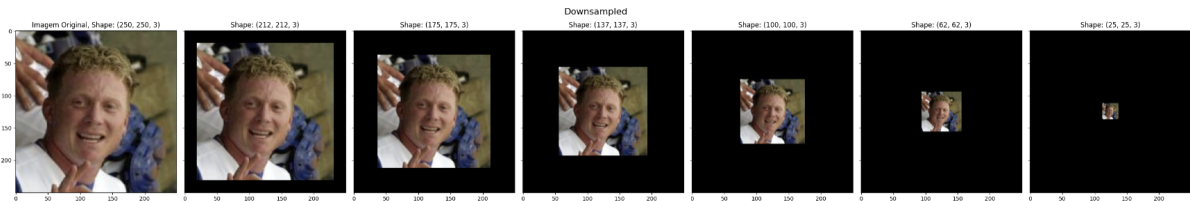


Figura 2.7: Redimensionado

### 2.5.4 Escurecimento

O Darkening, ou escurecimento, é uma técnica que consiste em diminuir o brilho global de uma imagem, resultando em uma aparência mais escura. Essa técnica pode ser útil para criar efeitos de sombras ou ajustar o contraste em determinadas áreas da imagem. [31]



Figura 2.8: Escurecimento

## 2.5.5 Desfoque de Movimento

O desfoque de movimento é uma degradação comumente encontrada em imagens capturadas com movimento da câmera ou objetos em movimento. Ele ocorre quando há um deslocamento do objeto durante o tempo de exposição da câmera. O efeito resultante é uma imagem borrada, com linhas alongadas na direção do movimento. [30]



Figura 2.9: Desfoque de Movimento

## 2.5.6 Brilho

O brilho é uma degradação que consiste em aumentar o brilho global da imagem, resultando em uma aparência mais clara. Essa técnica pode ser útil para criar efeitos de iluminação ou ajustar o contraste em determinadas áreas da imagem. [31]

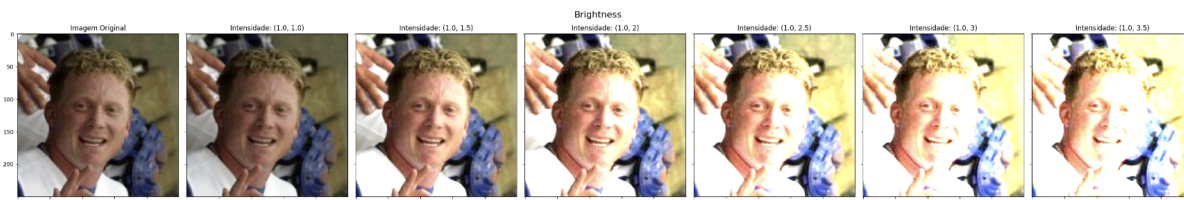


Figura 2.10: Brilho

## 2.5.7 Ruído de Sal e Pimenta

O ruído de sal e pimenta é um tipo de ruído aleatório que adiciona pixels pretos e brancos aleatoriamente na imagem. Esse tipo de ruído pode afetar a nitidez, a clareza e a qualidade geral da imagem. [29]



Figura 2.11: Ruído de Sal e Pimenta

## 2.5.8 Compressão JPEG

A compressão JPEG é um tipo de compressão com perdas, que reduz o tamanho da imagem removendo informações que não são perceptíveis ao olho humano. Esse tipo de compressão pode afetar a nitidez, a clareza e a qualidade geral da imagem. [30]

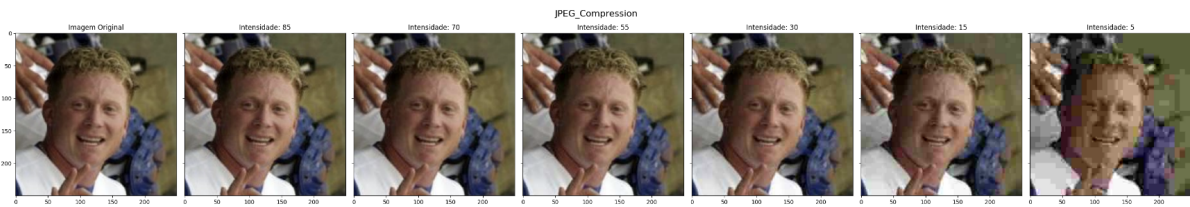


Figura 2.12: Compressão JPEG

# Capítulo 3

## Metodologia Proposta

A metodologia adotada neste estudo tem como objetivo fornecer uma abordagem sistemática para avaliar as degradações presentes em imagens faciais e sua interação com os modelos de Reconhecimento Facial que utilizam aprendizagem profunda. Isso permitirá entender o comportamento desses modelos em diferentes cenários de degradação, o que é essencial para avaliar sua robustez e eficácia em aplicações de uso geral. A estrutura básica dessa metodologia proposta é descrita na Figura 3.1.

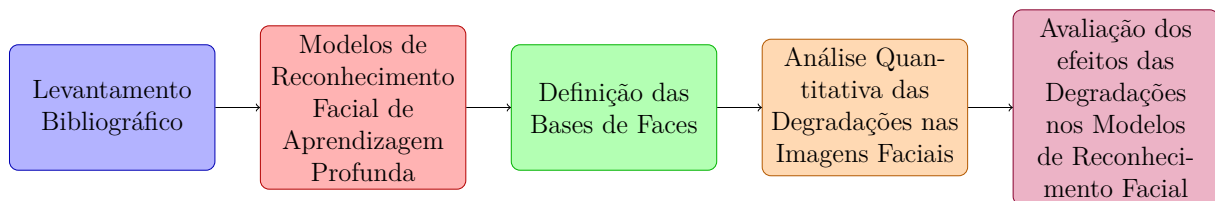


Figura 3.1: Fluxograma da Metodologia

### 3.1 Levantamento Bibliográfico das Degradações

Esta seção apresenta um levantamento bibliográfico das principais degradações que serão correlacionadas às imagens utilizadas em sistemas de reconhecimento facial. Inicialmente, o objetivo era abordar exclusivamente as degradações em imagens de faces, porém, ao longo da pesquisa, foram incluídos estudos que tratam de outras formas de degradação relevantes para a metodologia aplicada.

Para analisar o desempenho dos algoritmos de reconhecimento facial diante de imagens degradadas, serão utilizados oito modelos de reconhecimento facial: VGG [32], FaceNet [33], OpenFace [34], DeepFace [35], DeepID [36], ArcFace [37], Dlib [38] e SFace [27], conforme [6].



As degradações consideradas nos estudos incluem distorções como borramento de movimento, ruído, compressão e distorções de cor, conforme destacado por [39]. Essas degradações podem ocorrer em imagens capturadas por câmeras em movimento, em ambientes com pouca iluminação ou em condições de baixa qualidade de transmissão, por exemplo.

Ao analisar o impacto dessas degradações nos modelos de reconhecimento facial, os estudos visam contribuir para o desenvolvimento de técnicas e algoritmos mais robustos, capazes de lidar com imagens degradadas e garantir um desempenho confiável em cenários do mundo real.

Este levantamento bibliográfico serve como base para a compreensão das principais degradações que podem afetar as imagens utilizadas em sistemas de reconhecimento facial e proporciona uma visão geral dos modelos de reconhecimento facial utilizados nos estudos analisados.

## 3.2 Modelos de Reconhecimento Facial

Nesta seção, descreveremos os modelos de reconhecimento facial para análise dos estudos. A escolha desses modelos será baseada em sua relevância e desempenho comprovado em estudos anteriores.

Serão selecionados oito modelos de reconhecimento facial disponível e agrupados no projeto *Deep Face* [6]: VGG [32], FaceNet [33], OpenFace [34], DeepFace [35], DeepID [36], ArcFace [37], Dlib [38] e SFace [27]. Cada um desses modelos possui características e arquiteturas distintas, que oferecerão abordagens variadas para a tarefa de reconhecimento facial.

A escolha desses modelos será motivada por diversos fatores. Primeiramente, eles apresentarão alguns dos principais avanços no campo do reconhecimento facial utilizando aprendizado profundo. Cada modelo deverá possuir um conjunto específico de técnicas e arquiteturas que lhe conferem habilidades únicas na extração e representação de características faciais.

Além disso, esses modelos serão amplamente avaliados em termos de desempenho e robustez em cenários de degradação, incluindo o reconhecimento de faces em condições degradadas. Com o intuito de validar estudos anteriores demonstraram a eficácia desses modelos na presença de distorções, como borramento de movimento, ruído, compressão e distorções de cor, conforme mencionado na seção anterior.

Cada modelo selecionado apresenta vantagens e limitações específicas, que serão consideradas durante a análise dos resultados. Esperamos que essa diversidade de abordagens e características dos modelos contribua para uma compreensão mais abrangente do de-

sempenho do reconhecimento facial diante de imagens degradadas. É importante ressaltar que, embora nossa análise se concentre no desempenho desses modelos em relação às degradações nas imagens das faces, eles também são amplamente aplicados em outras áreas relacionadas, como reconhecimento de emoções, detecção de idade e gênero, entre outros.

No próximo estágio da metodologia, aplicaremos esses modelos aos conjuntos de dados contendo imagens degradadas e avaliaremos seu desempenho em termos de precisão, robustez e capacidade de generalização. Os resultados obtidos serão discutidos e analisados com o objetivo de fornecer uma perspectiva ampla sobre a eficácia dos modelos de reconhecimento facial em condições degradadas.

### 3.3 Bases de Faces

Nesta seção, será descrita a base de dados selecionada para a nossa pesquisa: a LFW (*Labeled Face in the Wild*). Explicaremos os motivos pelos quais escolhemos essa base específica e discutiremos sua relevância para a análise do desempenho dos modelos de reconhecimento facial em condições degradadas.

A LFW é uma das bases de dados mais amplamente utilizadas para avaliar algoritmos de reconhecimento facial em cenários do mundo real. Esta base contém uma grande variedade de imagens de faces coletadas a partir da internet, representando uma diversidade de pessoas, poses, expressões faciais, condições de iluminação e fundos.

A escolha da LFW como base de dados para nossa pesquisa foi motivada por alguns aspectos importantes. Primeiramente, sua natureza realista e diversificada permite a avaliação do desempenho dos modelos de reconhecimento facial em condições semelhantes às encontradas na vida cotidiana. Isso nos permite ter uma visão mais abrangente de como esses modelos se comportam diante de desafios relacionados a degradações nas imagens das faces.

Além disso, a LFW já foi amplamente utilizada em estudos anteriores, o que nos permite comparar nossos resultados com trabalhos relacionados e contribuir para a validação cruzada de descobertas na área. Sua popularidade e ampla aceitação também facilitam a comparação direta dos resultados obtidos com outros modelos e abordagens propostas pela comunidade científica.

A LFW fornece anotações de pares de imagens, indicando se duas imagens pertencem à mesma pessoa ou a pessoas diferentes. Essa informação é fundamental para a avaliação do desempenho dos modelos de reconhecimento facial, permitindo calcular métricas como taxa de acerto, taxa de falsos positivos e falsos negativos, entre outras.

Ao utilizar a LFW como base de dados, buscamos analisar como os modelos de reconhecimento facial selecionados se comportam em relação às degradações presentes nas

imagens de faces. As imagens da LFW serão submetidas a processos de degradação, como borrramento de movimento, adição de ruído, compressão e distorções de cor, a fim de simular condições desafiadoras. Avaliaremos o desempenho dos modelos em termos de precisão, robustez e capacidade de generalização nessas condições.

É importante ressaltar que, embora a LFW seja uma base de dados amplamente utilizada e ofereça uma diversidade significativa de faces e condições, ela apresenta algumas limitações, como um número limitado de imagens por indivíduo e possíveis vieses na seleção das imagens. No entanto, essas limitações não invalidam sua utilidade como um recurso de referência para avaliação e comparação de modelos de reconhecimento facial.

No próximo estágio da metodologia, aplicaremos os modelos de reconhecimento facial selecionados à base de dados LFW, tanto nas imagens originais quanto nas imagens degradadas. Os resultados obtidos serão analisados e comparados para fornecer uma visão abrangente do desempenho dos modelos em condições degradadas.

### 3.4 Quantificar Degradações

Nesta etapa da metodologia, realizaremos a quantificação das degradações presentes nas imagens de faces utilizando três arquivos .csv, nos quais serão registradas métricas para as análises quantitativas. Cada arquivo .csv corresponde a um tipo de comparação realizado entre as imagens: imagens degradadas **X** degradadas, imagens originais **X** degradadas e imagens originais **X** as mesmas imagens originais porém degradadas. Esses procedimentos foram repetidos para todas as degradações estudadas.

Essas métricas nos permitem quantificar o desempenho dos modelos de reconhecimento facial em relação às degradações nas imagens de faces. Podemos observar que diferentes modelos e algoritmos de detecção facial apresentam variações de desempenho, dependendo da intensidade da degradação. Por exemplo, no caso do Modelo VGG-Face com Detecção Dlib, podemos observar que, à medida que a intensidade da degradação aumenta, a acurácia diminui, assim como a precisão, *recall* e *f1-score*.

Essas informações quantitativas são fundamentais para avaliar a capacidade dos modelos em lidar com diferentes níveis de degradação e fornecem uma visão mais precisa de suas limitações e pontos fortes. Além disso, essas métricas permitem comparar o desempenho dos modelos em diferentes cenários de degradação e identificar possíveis estratégias para melhorar seu funcionamento.

Esse procedimento foi repetido para todas as degradações analisadas, como borrramento de movimento, adição de ruído, compressão e distorções de cor. Com base nessas métricas, poderemos avaliar o impacto de cada tipo de degradação no desempenho dos

modelos de reconhecimento facial e identificar quais modelos são mais robustos em condições desafiadoras.

No próximo estágio da metodologia, realizaremos a análise e interpretação dessas métricas, buscando identificar tendências e padrões nos resultados e fornecendo uma visão abrangente do desempenho dos modelos de reconhecimento facial frente às degradações nas imagens de faces.

### 3.5 Efeitos das Degradações

Nesta seção, será avaliado o impacto de diferentes tipos de degradação em cada modelo, visando identificar os modelos mais adequados para lidar com cada cenário. Para isso, foram conduzidos experimentos nos quais os dados de treinamento e teste foram submetidos a diferentes níveis de degradação.

Para cada tipo de degradação, os melhores modelos foram selecionados com base nas métricas de acurácia, precisão, *recall* e *fscore* obtidas nos experimentos. Essa abordagem permitiu identificar quais modelos são mais robustos em relação a determinados tipos de degradação e quais modelos são mais sensíveis a essas variações.

Os resultados dessa análise serão apresentados na seção de Resultados, juntamente com as métricas de desempenho dos modelos para cada tipo de degradação. Com base nessas informações, serão escolhidos os modelos que apresentaram melhor desempenho em cada tipo de degradação, os quais serão utilizados na fase de validação do sistema.

# Capítulo 4

## Resultados

Neste capítulo, apresentaremos uma sumarização dos resultados conforme os testes para as degradações utilizadas 2.5 e os respectivos modelos de reconhecimento facial utilizados 2.4. Os detalhes dos resultados estão mais detalhados em Anexo dos resultados. I

### 4.1 Base de imagens - *Labeled Faces in the Wild*

*Labeled Faces in the Wild* (LFW) é uma base de dados que contém 13.233 imagens de 5.749 pessoas, sendo que 1.680 dessas pessoas possuem duas ou mais imagens. Essas imagens foram coletadas através de pesquisas na Internet usando nomes de pessoas conhecidas, resultando em uma grande variedade de poses, expressões e condições de iluminação (PIE), bem como diferentes resoluções.

A base de dados LFW também inclui rótulos fornecidos por terceiros para atributos como gênero, etnia e outros. No entanto, é importante ressaltar que a base é extremamente desbalanceada em termos de representatividade de gênero e etnia. Isso significa que a quantidade de amostras disponíveis para certos grupos pode ser significativamente menor do que para outros, o que é interessante pois pode influenciar nos resultados de algoritmos de reconhecimento facial treinados com essa base.

### 4.2 Descrição dos Testes

Nesta seção, apresentaremos os resultados, agrupando as métricas em testes, em que para cada degradação e algoritmo de reconhecimento será apresentado os resultados das métricas obtidas conforme a seguinte combinação de testes:

- **Teste 1)** Imagem degradada x degradada até intensidade máxima;
- **Teste 2)** Imagem original x degradada até intensidade máxima;

- **Teste 3)** Imagem original x mesma imagem original até intensidade de degradação máxima;

Sendo que para  $S$  a face detectada na base e suas respectivas combinações conforme ilustrado abaixo:

$$\begin{aligned} S_{iD} &= S_{jD} \rightarrow \mathbf{T1} \\ S_{io} &= S_{jD} \rightarrow \mathbf{T2} \\ S_{io} &= S_{iD} \rightarrow \mathbf{T3} \end{aligned} \tag{4.1}$$

Conforme a seguinte ilustração:

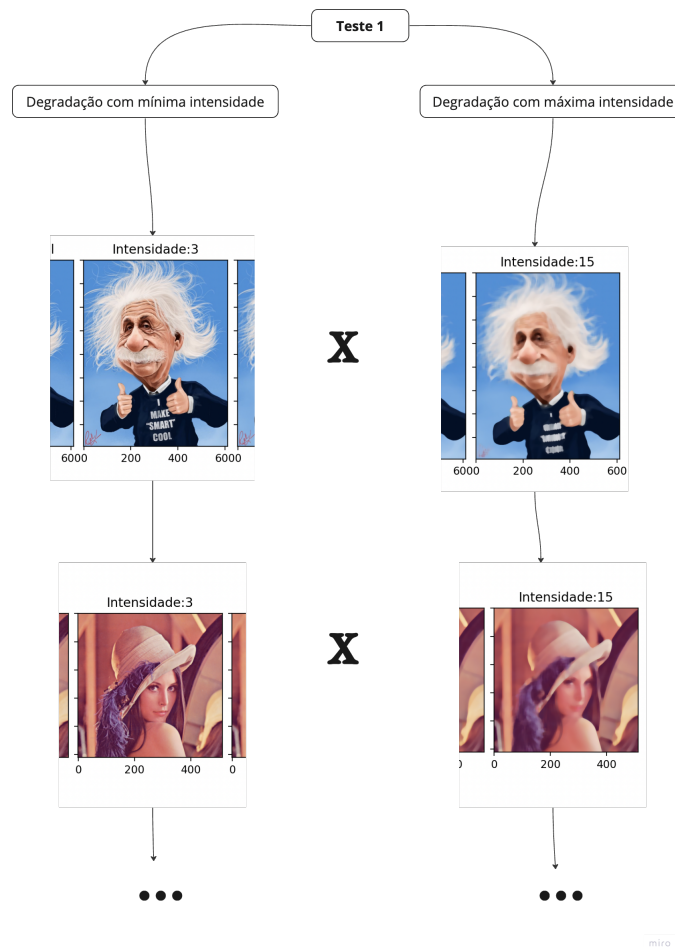


Figura 4.1: Ilustração de combinações para o Teste 1.

Desta forma para cada combinação de testes, distribuiu-se os resultados em uma tabela que demonstra os melhores e piores desempenhos conforme cada degradação consistindo em uma sumarização dos resultados obtidos.

## 4.3 Descrição dos Gráficos Elaborados

Nesta seção, apresentamos a descrição dos gráficos com os termos relacionados às métricas de desempenho utilizadas, estas fornecidas pelos modelos de reconhecimento facial selecionados, a saber:

### 4.3.1 Precision P+

A precision P+ (positiva), também conhecida como precisão positiva, é a proporção de instâncias corretamente classificadas como positivas (verdadeiro positivo) em relação ao total de instâncias classificadas como positivas (verdadeiro positivo + falso positivo). Essa métrica mede a habilidade do sistema em evitar falsos positivos, ou seja, em classificar corretamente as instâncias como positivas sem cometer erros de classificação positiva incorreta.

### 4.3.2 Precision P-

A precision P- (negativa), também conhecida como precisão negativa, é a proporção de instâncias corretamente classificadas como negativas (verdadeiro negativo) em relação ao total de instâncias classificadas como negativas (verdadeiro negativo + falso negativo). Essa métrica mede a habilidade do sistema em evitar falsos negativos, ou seja, em classificar corretamente as instâncias como negativas sem cometer erros de classificação negativa incorreta.

### 4.3.3 Recall P+

O recall P+ (positivo), também conhecido como taxa de verdadeiro positivo ou sensibilidade, é a proporção de instâncias corretamente classificadas como positivas (verdadeiro positivo) em relação ao total de instâncias que realmente são positivas (verdadeiro positivo + falso negativo). Essa métrica mede a habilidade do sistema em identificar corretamente as instâncias positivas, ou seja, em encontrar todas as instâncias positivas presentes no conjunto de dados.

### 4.3.4 Recall P-

O recall P- (negativo), também conhecido como taxa de verdadeiro negativo ou especificidade, é a proporção de instâncias corretamente classificadas como negativas (verdadeiro negativo) em relação ao total de instâncias que realmente são negativas (verdadeiro negativo + falso positivo). Essa métrica mede a habilidade do sistema em identificar

corretamente as instâncias negativas, ou seja, em encontrar todas as instâncias negativas presentes no conjunto de dados.

Essas métricas são amplamente utilizadas na avaliação de algoritmos de detecção e classificação, fornecendo informações importantes sobre a precisão e a capacidade de recall do modelo em diferentes cenários. É importante considerar tanto a precisão quanto o recall para avaliar o desempenho geral do algoritmo.

## 4.4 Teste 1 - Imagem degradada x degradada

Este primeiro conjunto de resultados apresenta os dados obtidos na primeira etapa dos experimentos, na qual foram analisadas imagens degradadas com diferentes intensidades, comparando-as com a base correspondente de imagens degradadas. Esses resultados ajudam a investigar como a degradação progressiva das imagens, até atingir a intensidade máxima, afeta ou não a qualidade e a capacidade de reconhecimento facial. Vale ressaltar que foram utilizados métodos de degradação combinados para gerar conjuntos de imagens com variação gradual na intensidade das degradações. Os resultados obtidos nessa etapa serão fundamentais para compreender o impacto das degradações nos algoritmos de reconhecimento e detecção, além de fornecer uma base de conhecimento empírico para o desenvolvimento de sistemas de aprendizado profundo.

Tabela 4.1: Tabela de testes 1

DEGRADAÇÃO		ArcFace			DeepFace	
Ruído Gaussiano	LM= 0.01 , LP = 0.15	M = 0.989 (dlib)	P = 0.028(mtcnn)	LM= 0.01 , LP = 0.15	M = 0.996 (dlib)	P = 0.571 (retinaface)
Borramento Gaussiano	LM = 3 , LP = 23	M = 0.990 (dlib)	P = 0.003 (dlib)	LM = 3 , LP = 23	M = 0.996 (dlib)	P = 0.495 (retinaface)
Redimensionado	LM = 70 , LP = 10	M = 0.990 (dlib)	P = 0.002 (dlib)	LM = 85 , LP = 25	M = 0.996 (dlib)	P = 0.713 (retinaface)
Escurecimento	LM = 0.9 , LP = 0.2	M = 0.990 (dlib)	P = 0.623 (mtcnn)	LM = 0.9 , LP = 0.2	M = 0.996 (dlib)	P = 0.325 (retinaface)
Desfoque de Movimento	LM = 3 , LP = 13	M = 0.990 (dlib)	P = 0.094 (mtcnn)	LM = 3 , LP = 13	M = 0.996 (dlib)	P = 0.472 (retinaface)
Brilho	LM = (1.0, 1.5), LP = (1.0, 3.5)	M = 0.991 (dlib)	P = 0.455(mtcnn)	LM = (1.0, 1.5), LP = (1.0, 3)	M = 0.996 (dlib)	P = 0.660 (retinaface)
Ruído de Sal e Pimenta	LM= 0.01 , LP = 0.15	M = 0.986 (dlib)	P = 0.029(retinaface)	LM= 0.01 , LP = 0.15	M = 0.996 (dlib)	P = 0.508 (mtcnn)
Compressão JPEG	LM = 55 , LP = 5	M = 0.990 (dlib)	P = 0.640(retinaface)	LM = 85 , LP = 15	M = 0.996 (dlib)	P = 0.705 (retinaface)
DEGRADAÇÃO		DeepID			Dlib	
Ruído Gaussiano	LM= 0.01 , LP = 0.15	M = 0.989 (dlib)	P = 0.968 (mtcnn)	LM= 0.01 , LP = 0.15	M = 0.999 (dlib)	P = 0.594 (retinaface)
Borramento Gaussiano	LM = 3 , LP = 23	M = 0.989 (dlib)	P = 0.944 (mtcnn)	LM = 3 , LP = 23	M = 0.999 (dlib)	P = 0.699 (mtcnn)
Redimensionado	LM = 10 , LP = 25	M = 0.997 (dlib)	P = 0.968 (mtcnn)	LM = 85 , LP = 10	M = 0.999 (dlib)	P = 0.692 (mtcnn)
Escurecimento	LM = 0.2 , LP = 0.9	M = 0.997 (retinaface)	P = 0.970 (mtcnn)	LM = 0.9 , LP = 0.2	M = 0.999 (dlib)	P = 0.957(mtcnn)
Desfoque de Movimento	LM = 3 , LP = 13	M = 0.989 (dlib)	P = 0.961 (mtcnn)	LM = 3 , LP = 13	M = 0.999 (dlib)	P = 0.740 (mtcnn)
Brilho	LM = (1.0, 1.0) , LP = (1.0, 3.5)	M = 0.989 (dlib)	P = 0.562 (retinaface)	LM = (1.0, 1.0) , LP = (1.0, 3.5)	M = 0.999 (dlib)	P = 0.572 (mtcnn)
Ruído de Sal e Pimenta	LM = 0.15 , LP = 0.15	M = 0.996 (retinaface)	P = 0.939 (mtcnn)	LM= 0.01 , LP = 0.15	M = 0.998 (dlib)	P = 0.486 (mtcnn)
Compressão JPEG	LM = 5 , LP = 15	M = 0.992 (dlib)	P = 0.969 (mtcnn)	LM = 85 , LP = 5	M = 0.999 (dlib)	P = 0.871 (retinaface)
DEGRADAÇÃO		Facenet512			OpenFace	
Ruído Gaussiano	LM= 0.01 , LP = 0.15	M = 0.998 (dlib)	P = 0.828 (dlib)	LM= 0.01 , LP = 0.15	M = 0.997 (dlib)	P = 0.994 (mtcnn)
Borramento Gaussiano	LM = 3 , LP = 23	M = 0.998 (dlib)	P = 0.911 (mtcnn)	LM = 3 , LP = 23	M = 0.997 (dlib)	P = 0.932 (mtcnn)
Redimensionado	LM = 85 , LP = 10	M = 0.998 (dlib)	P = 0.862 (dlib)	LM = 55 , LP = 10	M = 0.997 (dlib)	P = 0.990 (mtcnn)
Escurecimento	LM = 0.9 , LP = 0.2	M = 0.998 (dlib)	P = 0.925 (dlib)	LM = 0.9 , LP = 0.2	M = 0.997 (dlib)	P = 0.973 (retinaface)
Desfoque de Movimento	LM = 3 , LP = 13	M = 0.998 (dlib)	P = 0.980 (dlib)	LM = 3 , LP = 13	M = 0.997 (dlib)	P = 0.958 (mtcnn)
Brilho	LM = (1.0, 1.0) , LP = (1.0, 3.5)	M = 0.998 (dlib)	P = 0.943 (mtcnn)	LM = (1.0, 1.0) , LP = (1.0, 3.5)	M = 0.997 (dlib)	P = 0.978 (retinaface)
Ruído de Sal e Pimenta	LM= 0.01 , LP = 0.15	M = 0.998 (dlib)	P = 0.771 (dlib)	LM= 0.01 , LP = 0.15	M = 0.997 (dlib)	P = 0.992 (retinaface)
Compressão JPEG	LM = 85 , LP = 5	M = 0.998 (dlib)	P = 0.996 (retinaface)	LM = 5 , LP = 85	M = 0.997 (dlib)	P = 0.996 (retinaface)
DEGRADAÇÃO		SFace			VGG-Face	
Ruído Gaussiano	LM= 0.01 , LP = 0.15	M = 0.996 (retinaface)	P = 0.840 (dlib)	LM= 0.01 , LP = 0.15	M = 0.769 (dlib)	P = 0.048 (mtcnn)
Borramento Gaussiano	LM = 3 , LP = 23	M = 0.995 (retinaface)	P = 0.828 (dlib)	LM = 11 , LP = 23	M = 0.802 (dlib)	P = 0.045 (mtcnn)
Redimensionado	LM = 40 , LP = 10	M = 0.997 (retinaface)	P = 0.343(dlib)	LM = 40 , LP = 10	M = 0.796 (dlib)	P = 0.150 (mtcnn)
Escurecimento	LM = 0.9 , LP = 0.9	M = 0.996 (retinaface)	P = 0.916 (dlib)	LM = 0.9 , LP = 0.2	M = 0.770 (dlib)	P = 0.390 (retinaface)
Desfoque de Movimento	LM = 3 , LP = 13	M = 0.996 (retinaface)	P = 0.771 (dlib)	LM = 3 , LP = 13	M = 0.776 (dlib)	P = 0.137 (mtcnn)
Brilho	LM = (1.0, 1.0) , LP = (1.0, 3.5)	M = 0.996 (retinaface)	P = 0.622 (mtcnn)	LM = (1.0, 2) , LP = (1.0, 3.5)	M = 0.823 (dlib)	P = 0.351 (retinaface)
Ruído de Sal e Pimenta	LM= 0.01 , LP = 0.15	M = 0.996 (retinaface)	P = 0.615 (dlib)	LM= 0.01 , LP = 0.15	M = 0.727 (dlib)	P = 0.022 (mtcnn)
Compressão JPEG	LM = 15 , LP = 5	M = 0.996 (retinaface)	P = 0.848 (dlib)	LM = 70 , LP = 5	M = 0.778 (dlib)	P = 0.483 (mtcnn)

Desta forma, **LM** e **LP** representam a intensidade da degradação no melhor e pior caso, respectivamente. Por outro lado, **M** e **P** indicam o melhor e pior caso, respectivamente,



referindo-se ao algoritmo de detecção utilizado, em que a representação de cada destaque na Tabela são descritos como:

- **LM** = Intensidade da degradação do melhor caso;
- **LP** = Intensidade da degradação do pior caso;
- **M** = Melhor valor da acurácia (algoritmo de detecção aplicado no caso);
- **P** = Pior valor da acurácia (algoritmo de detecção aplicado no caso);
- Melhor valor para a pior intensidade;
- Pior valor para a pior intensidade;

Abaixo os gráficos que representam o melhor (fig.4.2) e pior desempenho (fig. 4.3)

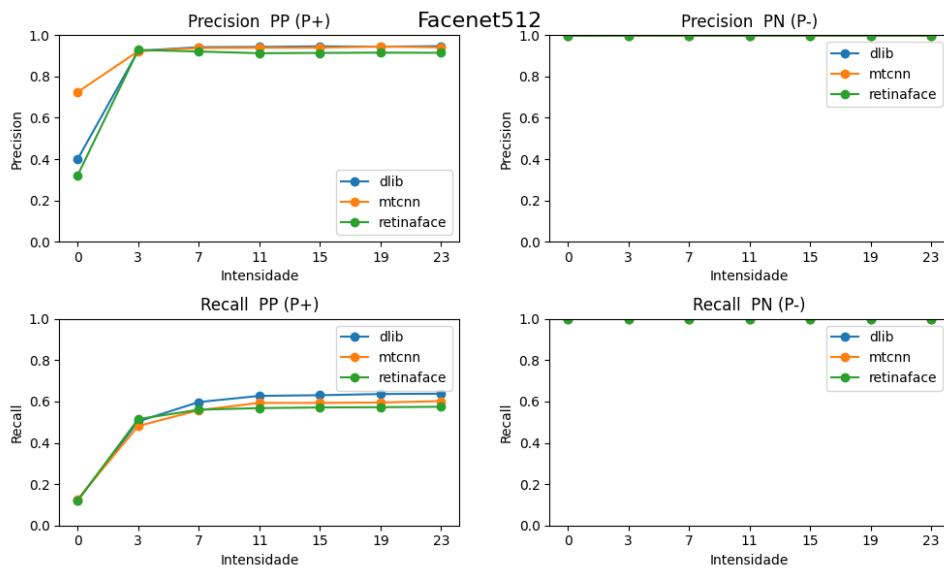


Figura 4.2: Melhor desempenho - Facenet512 - Compressão JPEG

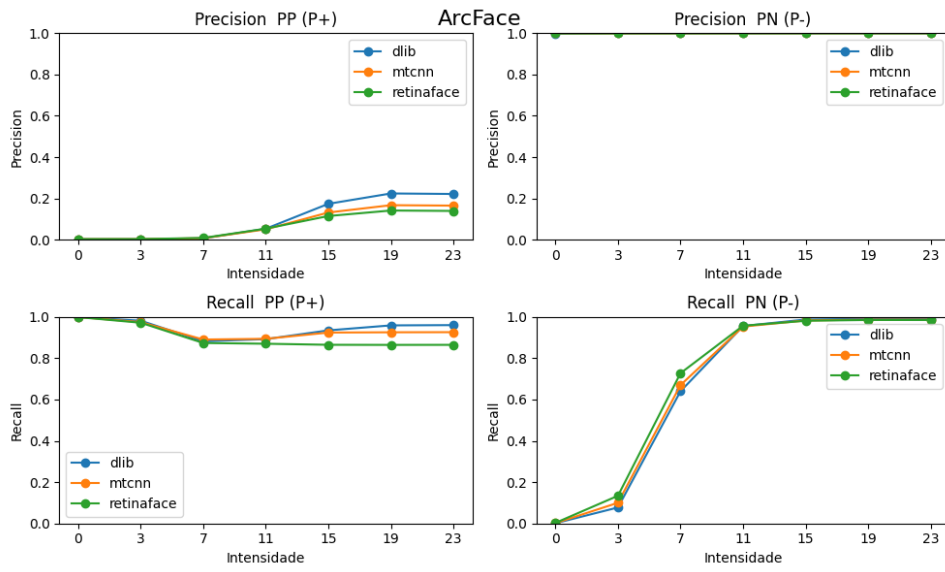


Figura 4.3: Pior desempenho - ArcFace Borramento Gaussiano

## 4.5 Teste 2 - Imagem original x degradada

Da mesma forma que o teste 1, o teste 2 consiste na análise comparativa entre imagens originais e suas versões degradadas. Nesta etapa dos experimentos, são avaliados os efeitos da degradação nas imagens originais, em contraste com a base correspondente de imagens degradadas. Os resultados obtidos nessa análise fornecem *insights* sobre a influência da degradação nas características faciais e na capacidade de reconhecimento e detecção. Nos diferentes métodos de degradação para gerar conjuntos de imagens com variação nas condições de degradação.

Tabela 4.2: Tabela de testes 2

DEGRADAÇÃO	ArcFace			DeepFace		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 0.990 (dlib)	P = 0.951 (retinaface)	LM= 0.23 , LP = 3	M = 0.997 (dlib)	P = 0.729 (retinaface)
Borramento Gaussiano	LM = 11 , LP = 23	M = 0.991 (dlib)	P = 0.952 (retinaface)	LM = 3 , LP = 23	M = 0.997 (dlib)	P = 0.735 (retinaface)
Redimensionado	LM = 40 , LP = 10	M = 0.991 (dlib)	P = 0.953 (retinaface)	LM = 10 , LP = 55	M = 0.963 (mtcnn)	P = 0.775 (mtcnn)
Escurecimento	LM = 0.0 , LP = 0.2	M = 0.990 (dlib)	P = 0.967 (retinaface)	LM = 0.2 , LP = 0.0	M = 0.997 (dlib)	P = 0.736 (retinaface)
Desfoque de Movimento	LM = 7 , LP = 13	M = 0.991 (dlib)	P = 0.956 (retinaface)	LM = 3 , LP = 5	M = 0.997 (dlib)	P = 0.716 (retinaface)
Brilho	LM = (1.0, 2) , LP = (1.0, 3.5)	M = 0.991 (dlib)	P = 0.970 (retinaface)	LM = (1.0, 3.5) , LP = (1.0, 1.0)	M = 0.997 (dlib)	P = 0.736 (retinaface)
Ruído de Sal e Pimenta	LM= 0.01 , LP = 0.15	M = 0.990 (dlib)	P = 0.952 (retinaface)	LM= 0.15 , LP = 0.1	M = 0.997 (dlib)	P = 0.735 (retinaface)
Compressão JPEG	LM = 0 , LP = 5	M = 0.990 (dlib)	P = 0.966 (retinaface)	LM = 0 , LP = 15	M = 0.996 (dlib)	P = 0.721 (retinaface)
DEGRADAÇÃO	DeepID			Dlib		
Ruído Gaussiano	LM= 0.15 , LP = 0.1	M = 0.990 (dlib)	P = 0.970 (mtcnn)	LM= 0.01 , LP = 0.15	M = 0.999 (dlib)	P = 0.953 (retinaface)
Borramento Gaussiano	LM = 0.0 , LP = 23	M = 0.989 (dlib)	P = 0.964 (mtcnn)	LM = 0.0 , LP = 0.0	M = 0.999 (dlib)	P = 0.964 (retinaface)
Redimensionado	LM = 10 , LP = 25	M = 0.997 (dlib)	P = 0.978 (mtcnn)	LM = 10 , LP = 25	M = 0.981 (mtcnn)	P = 0.968 (mtcnn)
Escurecimento	LM = 0.2 , LP = 0.0	M = 0.997 (dlib)	P = 0.970 (mtcnn)	LM = 0.0 , LP = 0.0	M = 0.999 (dlib)	P = 0.964 (retinaface)
Desfoque de Movimento	LM = 13 , LP = 13	M = 0.989 (dlib)	P = 0.968 (mtcnn)	LM = 0 , LP = 13	M = 0.999 (dlib)	P = 0.959 (mtcnn)
Brilho	LM = (1.0, 3.5) , LP = (1.0, 1.0)	M = 0.993 (dlib)	P = 0.970 (mtcnn)	LM = 0 , LP = (1.0, 1.0)	M = 0.999 (dlib)	P = 0.964 (retinaface)
Ruído de Sal e Pimenta	LM = 0.15 , LP = 0.15	M = 0.995 (retinaface)	P = 0.961 (mtcnn)	LM = 0.0 , LP = 0.07	M = 0.999 (dlib)	P = 0.961 (mtcnn)
Compressão JPEG	LM = 5 , LP = 15	M = 0.991 (dlib)	P = 0.970 (mtcnn)	LM = 0 , LP = 15	M = 0.999 (dlib)	P = 0.952 (retinaface)
DEGRADAÇÃO	Facenet512			OpenFace		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 0.998 (dlib)	P = 0.997 (dlib)	LM= 0.15 , LP = 0.0	M = 0.997 (dlib)	P = 0.996 (retinaface)
Borramento Gaussiano	LM = 0 , LP = 23	M = 0.998 (dlib)	P = 0.995 (mtcnn)	LM = 23 , LP = 7	M = 0.997 (dlib)	P = 0.996 (retinaface)
Redimensionado	LM = 0.0 , LP = 10	M = 0.998 (dlib)	P = 0.995 (retinaface)	LM = 10 , LP = 55	M = 0.997 (mtcnn)	P = 0.996 (mtcnn)
Escurecimento	LM = 0.0 , LP = 0.2	M = 0.998 (dlib)	P = 0.996 (retinaface)	LM = 0.2 , LP = 0.9	M = 0.997 (dlib)	P = 0.996 (retinaface)
Desfoque de Movimento	LM = 0 , LP = 13	M = 0.998 (dlib)	P = 0.996 (retinaface)	LM = 13 , LP = 5	M = 0.997 (dlib)	P = 0.995 (retinaface)
Brilho	LM = 0 , LP = (1.0, 3.5)	M = 0.998 (dlib)	P = 0.997 (retinaface)	LM = (1.0, 3.5) , LP = (1.0, 1.5)	M = 0.997 (dlib)	P = 0.995 (retinaface)
Ruído de Sal e Pimenta	LM= 0.0 , LP = 0.15	M = 0.998 (dlib)	P = 0.771 (dlib)	LM= 0.0 , LP = 0.0	M = 0.997 (dlib)	P = 0.996 (retinaface)
Compressão JPEG	LM = 0 , LP = 5	M = 0.998 (dlib)	P = 0.997 (retinaface)	LM = 5 , LP = 0	M = 0.997 (dlib)	P = 0.996 (retinaface)
DEGRADAÇÃO	SFace			VGG-Face		
Ruído Gaussiano	LM= 0.15 , LP = 0.0	M = 0.997 (retinaface)	P = 0.914 (dlib)	LM= 0.15 , LP = 0.1	M = 0.829 (dlib)	P = 0.559 (mtcnn)
Borramento Gaussiano	LM = 23 , LP = 3	M = 0.997 (mtcnn)	P = 0.913 (dlib)	LM = 23 , LP = 0	M = 0.975 (dlib)	P = 0.641 (mtcnn)
Redimensionado	LM = 25 , LP = 0.0	M = 0.997 (retinaface)	P = 0.914(dlib)	LM = 10 , LP = 55	M = 0.964 (mtcnn)	P = 0.636 (mtcnn)
Escurecimento	LM = 0.2 , LP = 0.0	M = 0.996 (retinaface)	P = 0.914 (dlib)	LM = 0.2 , LP = 0.4	M = 0.800 (dlib)	P = 0.628 (mtcnn)
Desfoque de Movimento	LM = 13 , LP = 3	M = 0.996 (mtcnn)	P = 0.910 (dlib)	LM = 13 , LP = 0	M = 0.914 (dlib)	P = 0.641 (mtcnn)
Brilho	LM = (1.0, 3.5) , LP = (1.0, 1.0)	M = 0.997 (mtcnn)	P = 0.914 (dlib)	LM = (1.0, 3.5) , LP = (1.0, 1.0)	M = 0.943 (dlib)	P = 0.641 (mtcnn)
Ruído de Sal e Pimenta	LM= 0.0 , LP = 0.0	M = 0.996 (retinaface)	P = 0.914 (dlib)	LM= 0.15 , LP = 0.07	M = 0.840 (dlib)	P = 0.586 (mtcnn)
Compressão JPEG	LM = 5 , LP = 0	M = 0.996 (retinaface)	P = 0.914 (dlib)	LM = 5 , LP = 15	M = 0.797 (dlib)	P = 0.637 (mtcnn)

- **LM** = Intensidade da degradação do melhor caso;
- **LP** = Intensidade da degradação do pior caso;
- **M** = Melhor valor da acurácia (algoritmo de detecção aplicado no caso);
- **P** = Pior valor da acurácia (algoritmo de detecção aplicado no caso);
- Melhor valor para a pior intensidade;
- Pior valor para a pior intensidade;

Abaixo os gráficos que representam o melhor (fig.4.4) e pior desempenho (fig. 4.5)

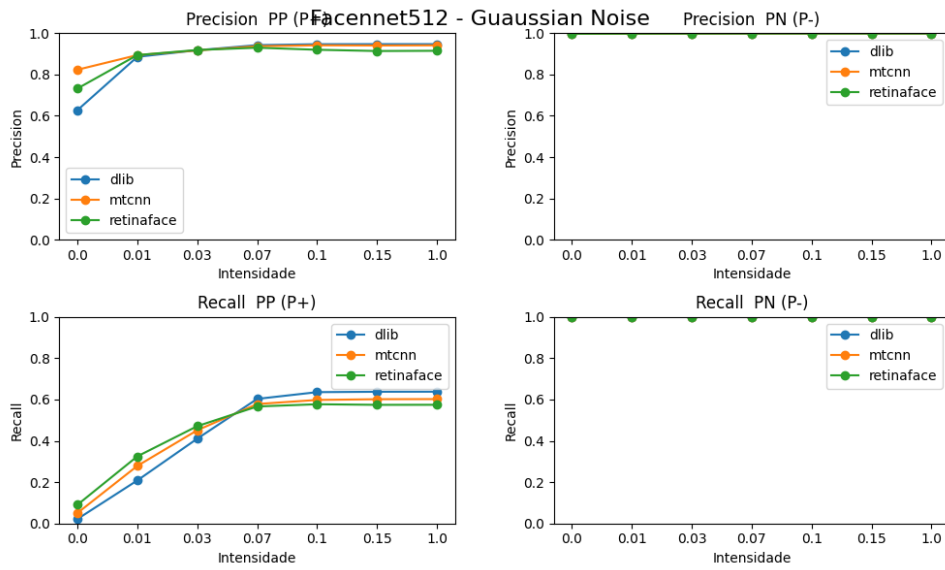


Figura 4.4: Melhor desempenho - Facenet512 Ruído Gaussiano

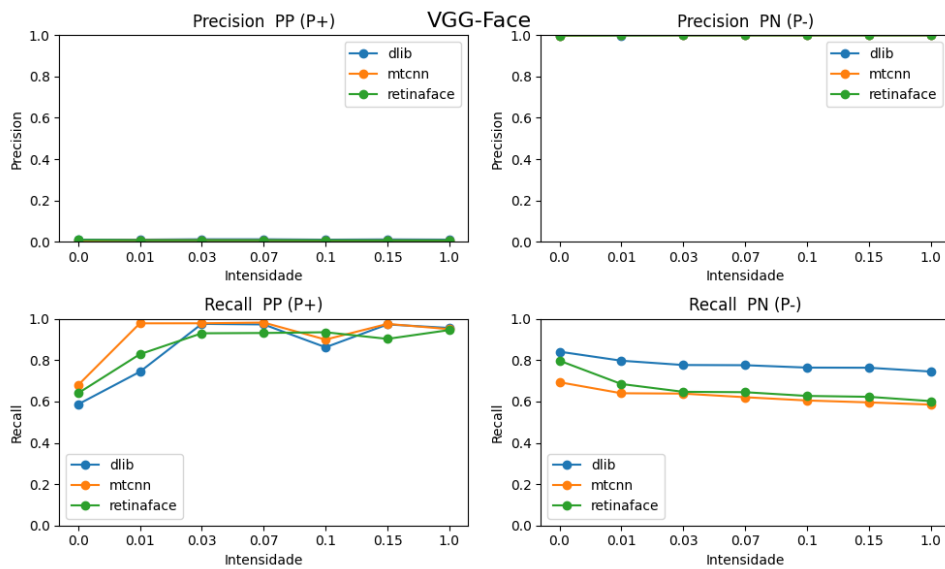


Figura 4.5: Pior desempenho - VGG-Face Ruído Sal e Pimenta

## 4.6 Teste 3 - Imagem original x mesma imagem original até intensidade máxima

Por fim, no teste 3, é realizada uma análise comparativa entre a imagem original e a mesma imagem original, porém submetida a um processo de degradação progressiva até atingir a intensidade máxima. Essa etapa dos experimentos visa investigar o impacto da degradação crescente na qualidade da imagem e na capacidade de reconhecimento facial. Utilizando diferentes métodos de degradação, são gerados conjuntos de imagens com variação gradual na intensidade das degradações aplicadas à imagem original.

Tabela 4.3: Tabela de testes 3

DEGRADAÇÃO	ArcFace			DeepFace		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 1.0 (dlib)	P = 0.107 (retinaface)	LM= 0.0 , LP = 0.15	M = 1.0 (retinaface)	P = 0.369 (dlib)
Borramento Gaussiano	LM = 0 , LP = 23	M = 1.0 (dlib)	P = 0.022 (dlib)	LM = 0 , LP = 23	M = 1.0 (mtcnn)	P = 0.164 (dlib)
Redimensionado	LM = 0 , LP = 10	M = 1.0 (dlib)	P = 0.012 (dlib)	LM = 0 , LP = 10	M = 1.0 (mtcnn)	P = 0.147 (dlib)
Escurecimento	LM = 0.0 , LP = 0.2	M = 1.0 (mtcnn)	P = 0.617 (dlib)	LM = 0.0 , LP = 0.2	M = 1.0 (dlib)	P = 0.051 (dlib)
Desfoque de Movimento	LM = 0 , LP = 13	M = 1.0 (mtcnn)	P = 0.197 (dlib)	LM = 0 , LP = 13	M = 1.0 (mtcnn)	P = 0.247 (dlib)
Brilho	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (retinaface)	P = 0.295 (mtcnn)	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (retinaface)	P = 0.118 (dlib)
Ruído de Sal e Pimenta	LM= 0 , LP = 0.15	M = 1.0 (mtcnn)	P = 0.102 (dlib)	LM= 0 , LP = 0.15	M = 1.0 (dlib)	P = 0.216 (retinaface)
Compressão JPEG	LM = 0 , LP = 5	M = 1.0 (mtcnn)	P = 0.615 (retinaface)	LM = 0 , LP = 15	M = 1.0 (dlib)	P = 0.583 (retinaface)
DEGRADAÇÃO	DeepID			Dlib		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 1.0 (dlib)	P = 0.579 (dlib)	LM= 0.0 , LP = 0.15	M = 1.0 (dlib)	P = 0.439 (dlib)
Borramento Gaussiano	LM = 0.0 , LP = 23	M = 1.0 (dlib)	P = 0.810 (mtcnn)	LM = 0.0 , LP = 23	M = 1.0 (retinaface)	P = 0.587 (mtcnn)
Redimensionado	LM = 0 , LP = 10	M = 1.0 (retinaface)	P = 0.004 (dlib)	LM = 0 , LP = 10	M = 1.0 (mtcnn)	P = 0.055 (mtcnn)
Escurecimento	LM = 0.2 , LP = 0.2	M = 7.556 (dlib)	P = 0 (mtcnn, retinaface)	LM = 0.0 , LP = 0.2	M = 1.0 (mtcnn)	P = 0.845 (dlib)
Desfoque de Movimento	LM = 0 , LP = 13	M = 1.0 (dlib)	P = 0.846 (dlib)	LM = 0 , LP = 13	M = 1.0 (dlib)	P = 0.602 (dlib)
Brilho	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (retinaface)	P = 0 (dlib)	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (dlib)	P = 0.241 (mtcnn)
Ruído de Sal e Pimenta	LM= 0 , LP = 0.15	M = 1.0 (mtcnn)	P = 0.135 (retinaface)	LM= 0 , LP = 0.15	M = 1.0 (retinaface)	P = 0.194 (retinaface)
Compressão JPEG	LM = 0 , LP = 5	M = 1.0 (dlib)	P = 0.485 (retinaface)	LM = 0 , LP = 5	M = 1.0 (retinaface)	P = 0.566 (retinaface)
DEGRADAÇÃO	Facenet512			OpenFace		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 1.0 (dlib)	P = 0.077 (dlib)	LM= 0.0 , LP = 0.15	M = 1.0 (retinaface)	P = 0.012 (dlib)
Borramento Gaussiano	LM = 0 , LP = 23	M = 1.0 (mtcnn)	P = 0.019 (mtcnn)	LM = 0 , LP = 23	M = 1.0 (dlib)	P = 0.003 (dlib)
Redimensionado	LM = 0.0 , LP = 10	M = 1.0 (retinaface)	P = 0.000 (dlib)	LM = 0.0 , LP = 10	M = 1.0 (mtcnn)	P = 0.001 (mtcnn)
Escurecimento	LM = 0.0 , LP = 0.2	M = 1.0 (mtcnn)	P = 0.196 (retinaface)	LM = 0.0 , LP = 0.2	M = 1.0 (mtcnn)	P = 0.001 (retinaface)
Desfoque de Movimento	LM = 0 , LP = 13	M = 1.0 (retinaface)	P = 0.036 (mtcnn)	LM = 0 , LP = 13	M = 1.0 (retinaface)	P = 0.005 (dlib)
Brilho	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (mtcnn)	P = 0.161 (mtcnn)	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (mtcnn)	P = 0.005 (mtcnn)
Ruído de Sal e Pimenta	LM= 0 , LP = 0.15	M = 1.0 (mtcnn)	P = 0.045 (dlib)	LM= 0 , LP = 0.15	M = 1.0 (dlib)	P = 0.001 (retinaface)
Compressão JPEG	LM = 0 , LP = 5	M = 1.0 (mtcnn)	P = 0.519 (retinaface)	LM = 0 , LP = 5	M = 1.0 (mtcnn)	P = 0.074 (retinaface)
DEGRADAÇÃO	SFace			VGG-Face		
Ruído Gaussiano	LM= 0.0 , LP = 0.15	M = 1.0 (retinaface)	P = 0.579 (dlib)	LM= 0.0 , LP = 0.15	M = 1.0 (mtcnn)	P = 0.543 (dlib)
Borramento Gaussiano	LM = 0 , LP = 23	M = 1.0 (retinaface)	P = 0.812 (dlib)	LM = 23 , LP = 0	M = 1.0 (dlib)	P = 0.171 (dlib)
Redimensionado	LM = 0 , LP = 10	M = 1.0 (dlib)	P = 0.010 (dlib)	LM = 0 , LP = 10	M = 1.0 (mtcnn)	P = 0.092 (mtcnn)
Escurecimento	LM = 0.0 , LP = 0.2	M = 1.0 (retinaface)	P = 0.776 (mtcnn)	LM = 0.0 , LP = 0.2	M = 1.0 (dlib)	P = 0.709 (dlib)
Desfoque de Movimento	LM = 0 , LP = 13	M = 1.0 (dlib)	P = 0.839 (retinaface)	LM = 0 , LP = 13	M = 1.0 (dlib)	P = 0.651 (dlib)
Brilho	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (mtcnn)	P = 0.111 (mtcnn)	LM = 0 , LP = (1.0, 3.5)	M = 1.0 (mtcnn)	P = 0.342 (mtcnn)
Ruído de Sal e Pimenta	LM= 0 , LP = 0.15	M = 1.0 (dlib)	P = 0.173 (retinaface)	LM= 0 , LP = 0.15	M = 1.0 (retinaface)	P = 0.534 (dlib)
Compressão JPEG	LM = 0 , LP = 5	M = 1.0 (retinaface)	P = 0.381 (retinaface)	LM = 0 , LP = 5	M = 1.0 (retinaface)	P = 0.894 (retinaface)

- **LM** = Intensidade da degradação do melhor caso;
- **LP** = Intensidade da degradação do pior caso;
- **M** = Melhor valor da acurácia (algoritmo de detecção aplicado no caso);
- **P** = Pior valor da acurácia (algoritmo de detecção aplicado no caso);
- Melhor valor para a pior intensidade;
- Pior valor para a pior intensidade;

Abaixo os gráficos que representam o melhor (fig.4.6) e pior desempenho (fig. 4.7)

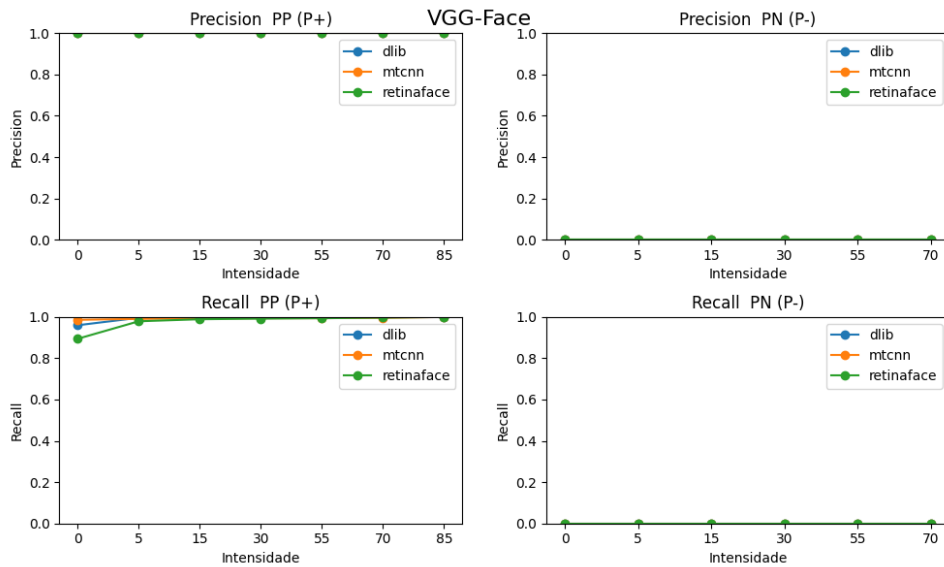


Figura 4.6: Melhor desempenho - VGG-Face Compressão JPEG

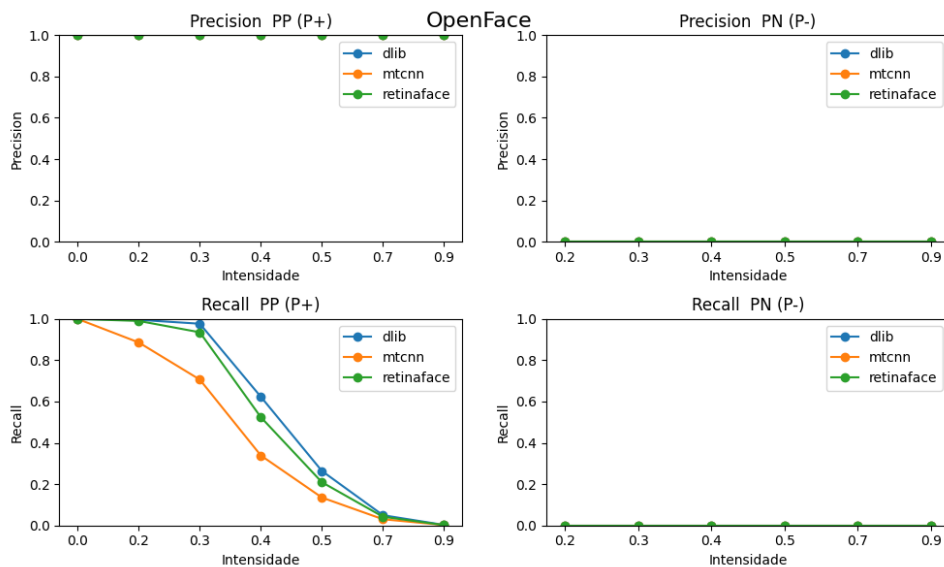


Figura 4.7: Pior desempenho - OpenFace Escurecimento

#### 4.6.1 Dos resultados

Analisando os resultados dos três testes realizados para avaliar o desempenho dos modelos de reconhecimento facial em diferentes degradações de imagens, é possível destacar algumas percepções importantes no contexto da área de reconhecimento facial. O Teste 1, que compara imagens degradadas entre si, mostrou que o modelo Arcface teve o pior desempenho com uma acurácia de apenas 0.002 na degradação por Borrachamento Gaussiano.

Por outro lado, os modelos Facenet512 e Openface apresentaram os melhores resultados, ambos com alta acurácia de 0.996 na degradação por Compressão JPEG.

No Teste 2, que avalia o desempenho dos modelos em imagens originais versus imagens degradadas, o modelo VGG-Face obteve o pior resultado, com uma acurácia de 0.586 na degradação por Ruído de Sal e Pimenta. Enquanto isso, o modelo Facenet512 alcançou a melhor acurácia com 0.997 na degradação por Ruído Gaussiano. Notavelmente, a diferença entre os melhores e piores resultados neste teste foi mais significativa do que no Teste 1.

Já no Teste 3, que compara imagens originais com elas mesmas, porém com intensidade máxima, o modelo Facenet512 teve o pior desempenho com acurácia de 0.000 na degradação por Redimensionado. Enquanto isso, o modelo VGG-Face obteve a melhor acurácia com 0.894 na degradação por Compressão JPEG.

De forma quantitativa, o Teste 2 apresentou o melhor resultado geral, já que os modelos alcançaram uma acurácia próxima a 1 (0.997) na degradação por Ruído Gaussiano. Por outro lado, o pior resultado geral foi obtido no Teste 3, onde o modelo Facenet512 teve uma acurácia de 0.000 na degradação por Redimensionado.

Esses resultados indicam que o desempenho dos modelos de reconhecimento facial pode variar significativamente dependendo do tipo de degradação aplicada às imagens. A degradação por Ruído Gaussiano se mostrou menos prejudicial ao reconhecimento facial, enquanto degradações como Borramento Gaussiano e Redimensionado tiveram um impacto mais negativo no desempenho dos modelos. Essas percepções destacam a importância de considerar as condições reais de aplicação e os tipos de degradações que podem ser encontrados em ambientes práticos ao projetar sistemas de reconhecimento facial robustos e precisos. Além disso, é fundamental selecionar modelos adequados para cada contexto específico, considerando suas características e comportamento diante de diferentes cenários de degradação.

# Capítulo 5

## Conclusões

Neste trabalho, foram abordados diversos aspectos relacionados à degradação de imagens faciais e seu impacto no desempenho de modelos de reconhecimento facial. Foram apresentados os principais modelos utilizados nesse contexto, bem como os métodos de degradação de imagens faciais empregados. Além disso, discutiu-se a importância de métodos de avaliação de desempenho para mensurar a eficácia desses modelos diante das imagens degradadas.

Conclui-se que a degradação de imagens faciais pode ter um impacto significativo no desempenho dos modelos de reconhecimento facial prejudicando sua capacidade de realizar corretamente as tarefas de identificação e autenticação, embora, como foi observado, o nível de intensidade não esteve diretamente relacionada com a performance dos algoritmos. Nesse sentido, é fundamental desenvolver técnicas e algoritmos robustos que possam lidar com essas degradações e manter um desempenho satisfatório.

O estudo dos efeitos da degradação de imagens faciais é relevante para aplicações em segurança, autenticação e identificação biométrica e não apenas para o campo do reconhecimento facial. A compreensão dos desafios e limitações associados a essas degradações pode contribuir para o desenvolvimento de soluções mais robustas e confiáveis, tornando assim os resultados atingidos nesse estudo, significantes para o avanço no campo.

É importante ressaltar que este trabalho é uma parte dos estudos e pesquisas feita pelo grupo de reconhecimento facial e existem várias áreas de pesquisa em aberto que podem ser exploradas para aprimorar o desempenho dos modelos de reconhecimento facial frente às imagens degradadas. Espera-se que futuros estudos e avanços nessa área possam proporcionar soluções mais eficazes e confiáveis para o reconhecimento facial em condições adversas.

Ao longo deste estudo, realizamos uma análise comparativa das métricas dos algoritmos de detecção facial em dois aspectos: as características faciais presentes no campo esquerdo da face e ao campo direito da face. Também consideramos diferentes intensida-



des de detecção para avaliar a sensibilidade dos algoritmos a variações na detecção facial, o que se fez evidente o impacto dessa análise nos resultados.

Portanto, com base nos resultados obtidos e melhor detalhado no Anexo I, podemos destacar algumas conclusões relevantes. Em relação à desempenho geral, observamos que o algoritmo Dlib apresentou um desempenho ligeiramente superior em comparação com os demais algoritmos de reconhecimento visto que observou-se um padrão mais normalizados, sem o mesmo está na classificação dos piores ou melhores. E em grande parte, sua combinação com algoritmo de detecção dlib, esse desempenho se torna ainda mais presente. Além do mais esse resultado sugere que o FaceNet512 e OpenFace possuem uma capacidade de detecção mais precisa em uma ampla gama de intensidades estando na classificação dos melhores resultados com valores superior a média. No entanto, é válido mencionar que todos os algoritmos alcançaram valores de precisão significativos, o que indica uma boa capacidade de detecção em geral.

É fundamental ressaltar que a escolha do algoritmo de detecção facial deve levar em consideração outros fatores além das métricas. Outros aspectos, como a velocidade de processamento, o consumo de recursos computacionais e a disponibilidade de implementações, também devem ser considerados ao selecionar o algoritmo mais adequado para uma aplicação específica. Desta forma, é mais evidente essa influência no desempenho dos algoritmos nas degradações de Brilho e Escurecimento visto que essas degradações apresentam um aspecto de intensidade que muito diferem das demais degradações, conforme

Em suma, este estudo nos forneceu *insights* valiosos sobre o desempenho dos algoritmos de reconhecimento e detecção facial. Os resultados obtidos destacam as diferenças e semelhanças entre esses algoritmos em diferentes cenários.

No entanto, é necessário mencionar que este estudo teve algumas limitações. A avaliação foi baseada em um conjunto de dados específico e pode não representar todas as possíveis variações e cenários encontrados no mundo real. Além disso, apenas três algoritmos de detecção (dlib,mtcnn,retinaface) foram considerados, e outros algoritmos de detecção facial também podem ser relevantes e dignos de investigação futura.

Em conclusão, este estudo proporciona um maior entendimento sobre a acurácia dos algoritmos de reconhecimento facial. Os resultados obtidos fornecem informações valiosas para pesquisadores e profissionais da área, permitindo uma avaliação mais precisa das capacidades e limitações desses algoritmos. Com base nessas descobertas, é possível tomar decisões mais embasadas ao selecionar e implementar algoritmos de detecção facial em aplicações do mundo real. Dessa forma, o estudo contribui sobremaneira para o avanço e aprimoramento contínuo da tecnologia de reconhecimento facial.

# Referências

- [1] Barreto, Saulo Cardoso, Jorge Albuquerque Lambert e Flávio de Barros Vidal: *Using synthetic images for deep learning recognition process on automatic license plate recognition*. Em Carrasco-Ochoa, Jesús Ariel, José Francisco Martínez-Trinidad, José Arturo Olvera-López e Joaquín Salas (editores): *Pattern Recognition*, páginas 115–126, Cham, 2019. Springer International Publishing, ISBN 978-3-030-21077-9. 1
- [2] Karahan, Samil, Merve Kilinc Yildirim, Kadir Kirtac, Ferhat Sukru Rende, Gultekin Butun e Hazim Kemal Ekenel: *How image degradations affect deep cnn-based face recognition?* Em *2016 International Conference of the Biometrics Special Interest Group (BIOSIG)*, páginas 1–5, 2016. 1
- [3] Grm, Klemen, Vitomir Štruc, Anais Artiges, Matthieu Caron e Hazım K. Ekenel: *Strengths and weaknesses of deep learning models for face recognition against image degradations*. *IET Biometrics*, 7(1):81–89, 2018. <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/iet-bmt.2017.0083>. 1
- [4] Rodríguez, Joaquín Salas, Flavio De Barros Vidal e Francisco Martinez-Trinidad: *Deep learning: Current state*. *IEEE Latin America Transactions*, 17(12):1925–1945, 2019. 2
- [5] Bicalho, Gustavo Carneiro, Marcelo Cobias Amorim Alves, Lucas Faria Porto, Carlos Eduardo Palhares Machado e Flavio de Barros Vidal: *Solving the face growth problem in the biometric face recognition using photo-anthropometric ratios by iris normalization*. Em *2018 International Workshop on Biometrics and Forensics (IWBF)*, páginas 1–6, 2018. 2
- [6] Serengil, Sefik Ilkin e Alper Ozpinar: *Lightface: A hybrid deep face recognition framework*. Em *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, páginas 23–27. IEEE, 2020. <https://doi.org/10.1109/ASYU50717.2020.9259802>. 2, 20, 21
- [7] Jain, Anil K., Brendan Klare e Unsang Park: *Face recognition: Some challenges in forensics*. Em *2011 IEEE International Conference on Automatic Face Gesture Recognition (FG)*, páginas 726–733, 2011. 5
- [8] Oliveira, Maria Olívia Domingos Rezio, Janaina Paiva Curi, Rosane Pérez Baldasso e Thiago Leite Beaini: *Reconhecimento facial na prática forense: Uma análise dos documentos disponibilizados pelo fiswg*. *Revista Brasileira de Odontologia Legal*, 9(1), 2022. 6

- [9] Chihaoui, Mejda, Akram Elkefi, Wajdi Bellil e Chokri Ben Amar: *A survey of 2d face recognition techniques*. Computers, 5(4):21, 2016. 6
- [10] Du, Hang, Hailin Shi, Dan Zeng, Xiao Ping Zhang e Tao Mei: *The elements of end-to-end deep face recognition: A survey of recent advances*. ACM Computing Surveys (CSUR), 54(10s):1–42, 2022. 6
- [11] Turk, Matthew e Alex Pentland: *Eigenfaces for Recognition*. Journal of Cognitive Neuroscience, 3(1):71–86, janeiro 1991, ISSN 0898-929X. <https://doi.org/10.1162/jocn.1991.3.1.71>. 7
- [12] Krizhevsky, Alex, Ilya Sutskever e Geoffrey E Hinton: *Imagenet classification with deep convolutional neural networks*. Communications of the ACM, 60(6):84–90, 2017. 7
- [13] Chen, Qiang, Feiping Nie, Weizhong Yu e Xuelong Li:  *$\ell_{2,p}$ -norm and mahalanobis distance based robust fuzzy c-means*. IEEE Transactions on Fuzzy Systems, páginas 1–13, 2023. 9
- [14] Souza Teixeira, Queila Valentim Martins de: *Análise de imagens de folhas de tabaco afetadas por cvc via técnicas de ranqueamento e re-ranqueamento*. 10
- [15] Santos, Erick Ramos dos, Patrick Marques Ciarelli e Jorge Leonid Aching Samatelo: *Constellation loss em modelos de reconhecimento facial*. Em *Simpósio Brasileiro de Automação Inteligente-SBAI*, volume 1, 2021. 10
- [16] Santos, Anderson Rodrigo dos: *Identificação de faces humanas através de PCA-LDA e redes neurais SOM*. Tese de Doutorado, Universidade de São Paulo, 2005. 11
- [17] Ishida, Tales Hiro Cardoso: *Sistema de controle de acesso por reconhecimento facial utilizando sistemas embarcados*. 2023. 11
- [18] Faveri, Eduardo Martins: *Estudo da eficácia dos algoritmos de reconhecimento facial*. 2023. 12
- [19] Ma, Chaoqun, Xiaoyu Jiang, Jing Liu e Liupeng Li: *A novel feed-forward neural network-based method for fast hologram generation*. Optics Communications, 530:129162, 2023. 12
- [20] Zisserman, Andrew e Stephen Maybank: *A case against epipolar geometry*. Em Mundy, J., A. Zisserman e D. Forsyth (editores): *Applications of Invariance in Computer Vision*. Springer-Verlag, 1994. 13
- [21] Balaban, Stephen: *Deep learning and face recognition: the state of the art*. Biometric and surveillance technology for human and activity identification XII, 9457:68–75, 2015. 13
- [22] Baltrušaitis, Tadas, Peter Robinson e Louis Philippe Morency: *Openface: An open source facial behavior analysis toolkit*. Em *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, páginas 1–10, 2016. 13

- [23] Zhou, Erjin, Zhimin Cao e Qi Yin: *Naive-deep face recognition: Touching the limit of lfw benchmark or not?* arXiv preprint arXiv:1501.04690, 2015. 14
- [24] Sun, Yi, Ding Liang, Xiaogang Wang e Xiaoou Tang: *Deepid3: Face recognition with very deep neural networks*, 2015. 14
- [25] Deng, Jiankang, Jia Guo, Niannan Xue e Stefanos Zafeiriou: *Arcface: Additive angular margin loss for deep face recognition*. Em *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, páginas 4690–4699, 2019. 14
- [26] Suwarno, Suwarno e Kevin Kevin: *Analysis of face recognition algorithm: Dlib and opencv*. *Journal of Informatics and Telecommunication Engineering*, 4(1):173–184, 2020. 14
- [27] Zhong, Yaoyao, Weihong Deng, Jiani Hu, Dongyue Zhao, Xian Li e Dongchao Wen: *Sface: Sigmoid-constrained hypersphere loss for robust face recognition*. *IEEE Transactions on Image Processing*, 30:2587–2598, 2021. 15, 20, 21
- [28] Deepthi, RS e S Sankaraiah: *Implementation of mobile platform using qt and opencv for image processing applications*. Em *2011 IEEE Conference on Open Systems*, páginas 284–289. IEEE, 2011. 15
- [29] Pratt, William K: *Digital image processing: PIKS Scientific inside*, volume 4. Wiley Online Library, 2007. 16, 18
- [30] Gonzalez, RC, RE Woods e SL Eddins: *Digital image processing using matlab” gate-mark publishing*, 2009. 17, 18, 19
- [31] Burger, Wilhelm e Mark J Burge: *Digital image processing: an algorithmic introduction using Java*. Springer, 2016. 17, 18
- [32] Parkhi, Omkar M, Andrea Vedaldi e Andrew Zisserman: *Deep face recognition*. 2015. 20, 21
- [33] Schroff, Florian, Dmitry Kalenichenko e James Philbin: *Facenet: A unified embedding for face recognition and clustering*. CoRR, abs/1503.03832, 2015. <http://arxiv.org/abs/1503.03832>. 20, 21
- [34] Amos, Brandon, Bartosz Ludwiczuk e Mahadev Satyanarayanan: *Openface: A general-purpose face recognition library with mobile applications*. Relatório Técnico, CMU-CS-16-118, CMU School of Computer Science, 2016. 20, 21
- [35] Taigman, Yaniv, Ming Yang, Marc’Aurelio Ranzato e Lior Wolf: *Deepface: Closing the gap to human-level performance in face verification*. Em *Proceedings of the IEEE conference on computer vision and pattern recognition*, páginas 1701–1708, 2014. 20, 21
- [36] Sun, Yi, Xiaogang Wang e Xiaoou Tang: *Deep learning face representation from predicting 10,000 classes*. Em *Proceedings of the IEEE conference on computer vision and pattern recognition*, páginas 1891–1898, 2014. 20, 21

- [37] Leondgarse: *Keras insightface*. [https://github.com/leondgarse/Keras\\_insightface](https://github.com/leondgarse/Keras_insightface), 2022. 20, 21
- [38] King, Davis E.: *Dlib-ml: A machine learning toolkit*. *Journal of Machine Learning Research*, 10:1755–1758, 2009. 20, 21
- [39] Karahan, Samil, Merve Kilinc Yildirim, Kadir Kirtac, Ferhat Sukru Rende, Gultekin Butun e Hazim Kemal Ekenel: *How image degradations affect deep cnn-based face recognition?* Em *2016 international conference of the biometrics special interest group (BIOSIG)*, páginas 1–5. IEEE, 2016. 21

# Anexo I

## Resultados Completos

### Anexo dos resultados

Nesta seção, serão apresentados os resultados dos estudos de forma mais detalhada, fornecendo uma análise aprofundada que sustenta as conclusões discutidas anteriormente. Além das métricas de acurácia, serão exploradas duas métricas adicionais que desempenham um papel crucial na avaliação do desempenho geral de cada algoritmo: precisão e taxa de recuperação.

A precisão mede a proporção de resultados corretos entre as detecções realizadas pelo algoritmo. Ela fornece uma compreensão sobre a capacidade do algoritmo de evitar falsos positivos e garantir que apenas as faces corretas sejam identificadas com precisão.

Por outro lado, a taxa de recuperação é uma métrica que indica a capacidade do algoritmo de detectar corretamente todas as faces presentes em uma imagem. Ela mede a proporção de faces corretamente detectadas em relação ao número total de faces presentes.

A análise dessas métricas adicionais nos ajudará a obter uma compreensão mais completa do desempenho de cada algoritmo de detecção facial, permitindo-nos avaliar não apenas a acurácia geral, mas também a precisão e a taxa de recuperação. Essa análise detalhada contribuirá para uma melhor avaliação e comparação dos algoritmos em termos de sua eficácia e capacidade de atender aos requisitos específicos de aplicação. Consulte a sessão Legenda dos Gráficos 4.3

### I.1 No contexto do Teste 1

#### Imagem degradada x degradada

Nas métricas geradas dessa combinação, observaram-se os seguintes comportamentos acerca das degradações e os algoritmos de reconhecimento com os piores desempenhos.

### I.1.1 Borramento Gaussiano

Dentre os resultados das nove degradações analisadas e estudadas nesse projeto o ruído gaussiano se destacou devido certos aspectos interessantes de desempenho dos modelos os quais serão mostrados nessa sessão.

#### ArcFace

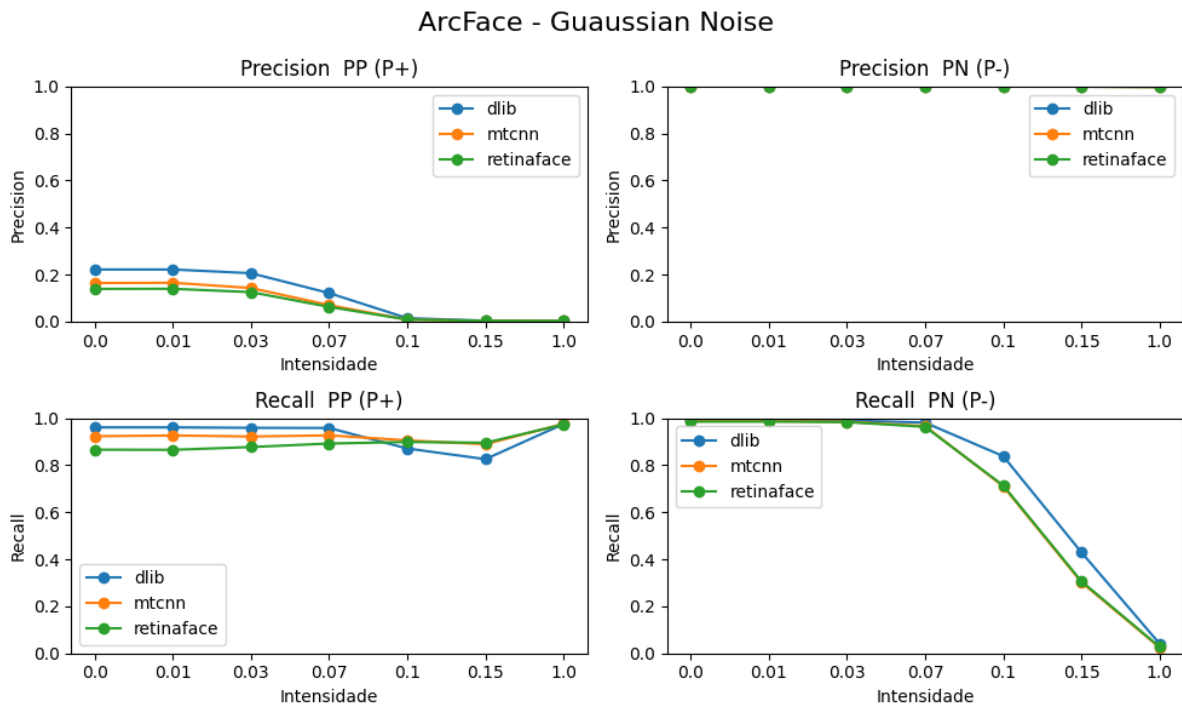


Figura I.1: Borramento Gaussiano - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

Observa-se que, para a métrica de precisão, à medida que a intensidade do borramento gaussiano aumenta, a precisão do algoritmo ArcFace também diminui. Isso indica que o borramento gaussiano tem um impacto negativo na capacidade do algoritmo de reconhecer corretamente as faces nas imagens. Quanto maior a intensidade do borramento, maior é a perda de precisão do algoritmo.

No entanto, é importante ressaltar que, mesmo com o aumento da intensidade do borramento, o algoritmo ainda é capaz de realizar o reconhecimento facial, embora com uma taxa de precisão reduzida. Isso sugere que o ArcFace é relativamente robusto em

lidar com a degradação por borramento gaussiano, mas sua eficácia é comprometida à medida que a degradação aumenta.

Com base nas informações fornecidas, ao analisar as curvas de taxa de recuperação para a degradação por borramento gaussiano, observa-se que a curva se mantém constante até atingir um determinado valor de intensidade de borramento. Nesse ponto, ocorre uma anomalia na curva P+ e uma queda na curva P-.

Isso indica que, até certo limite de intensidade de borramento, o algoritmo de reconhecimento facial apresenta uma taxa de recuperação constante, ou seja, é capaz de identificar corretamente as faces nas imagens degradadas. No entanto, após esse limite, ocorre uma anomalia na curva P+, indicando uma redução na capacidade do algoritmo de identificar corretamente as faces positivas, ou seja, aquelas que devem ser reconhecidas.

Além disso, há uma queda na curva P-, indicando uma diminuição na capacidade do algoritmo de rejeitar corretamente as faces negativas, ou seja, aquelas que não devem ser reconhecidas.

Essa anomalia e queda nas curvas de recall podem ser atribuídas ao aumento da intensidade do borramento gaussiano, que torna as características faciais menos distintas e mais difíceis de serem identificadas pelo algoritmo de reconhecimento facial.



## DeepFace

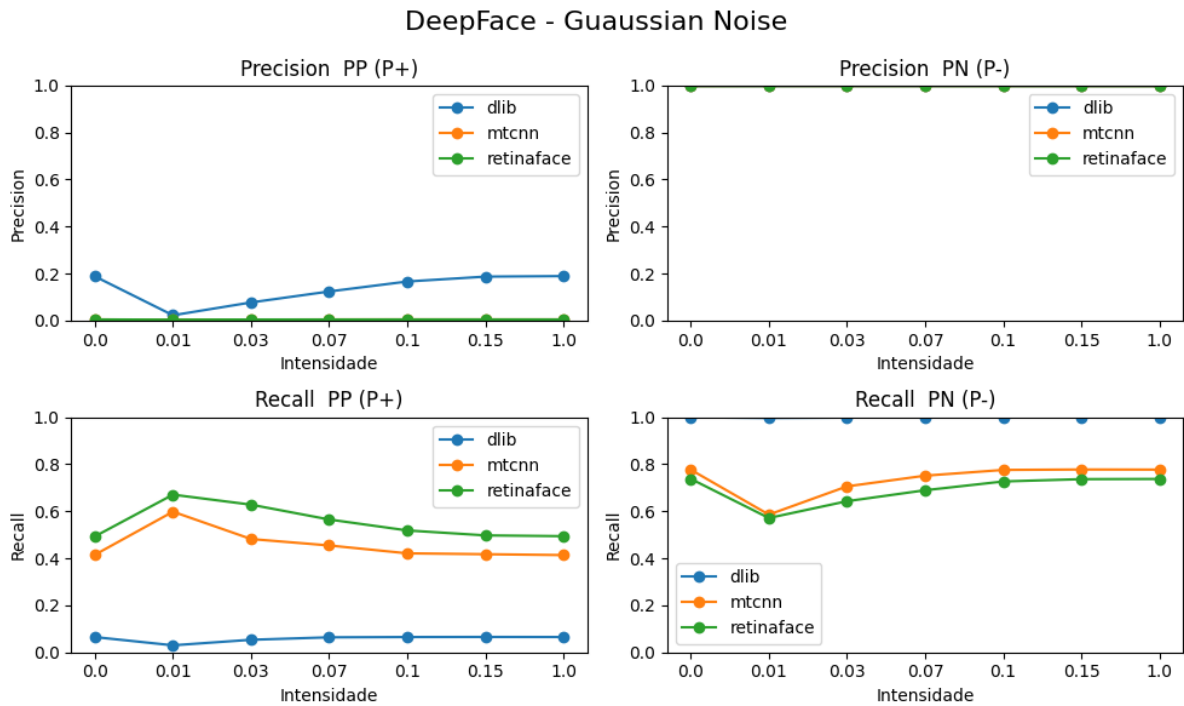


Figura I.2: Borramento Gaussiano - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original. No contexto da degradação por borramento gaussiano, o algoritmo DeepFace demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do borramento aumenta.

## Dlib

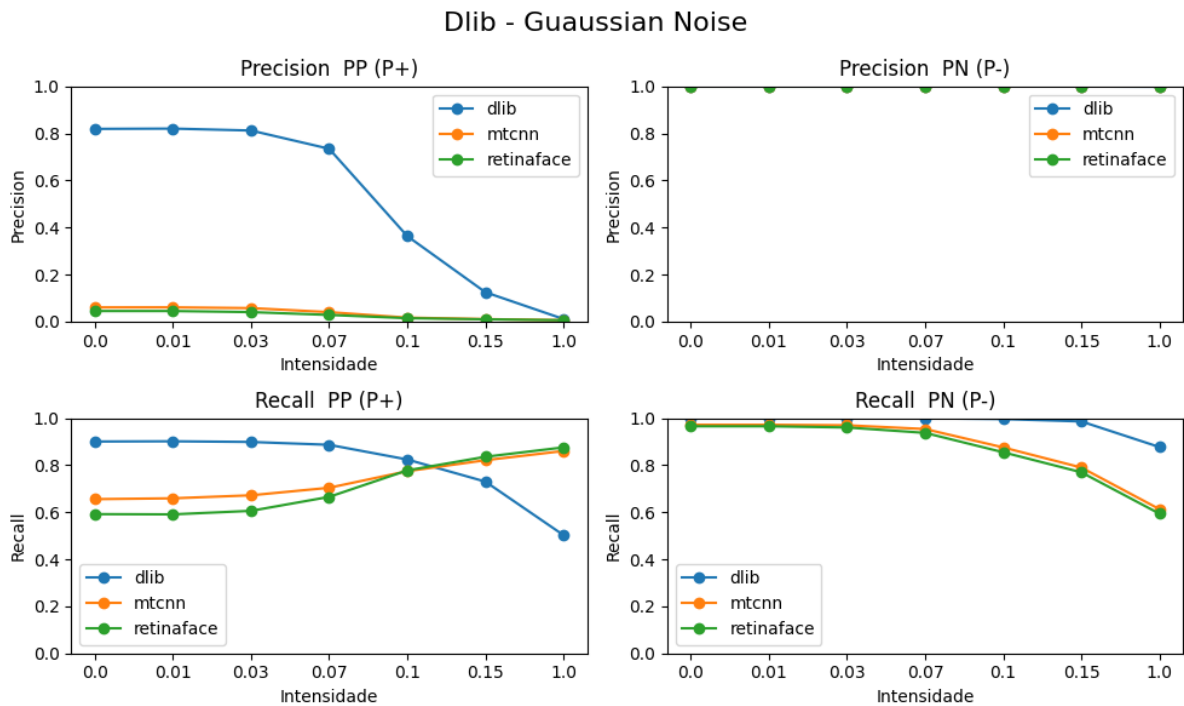


Figura I.3: Borramento Gaussiano - Dlib

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo Dlib em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original. No contexto da degradação por borramento gaussiano, o algoritmo Dlib demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do borramento aumenta.

## Facenet512

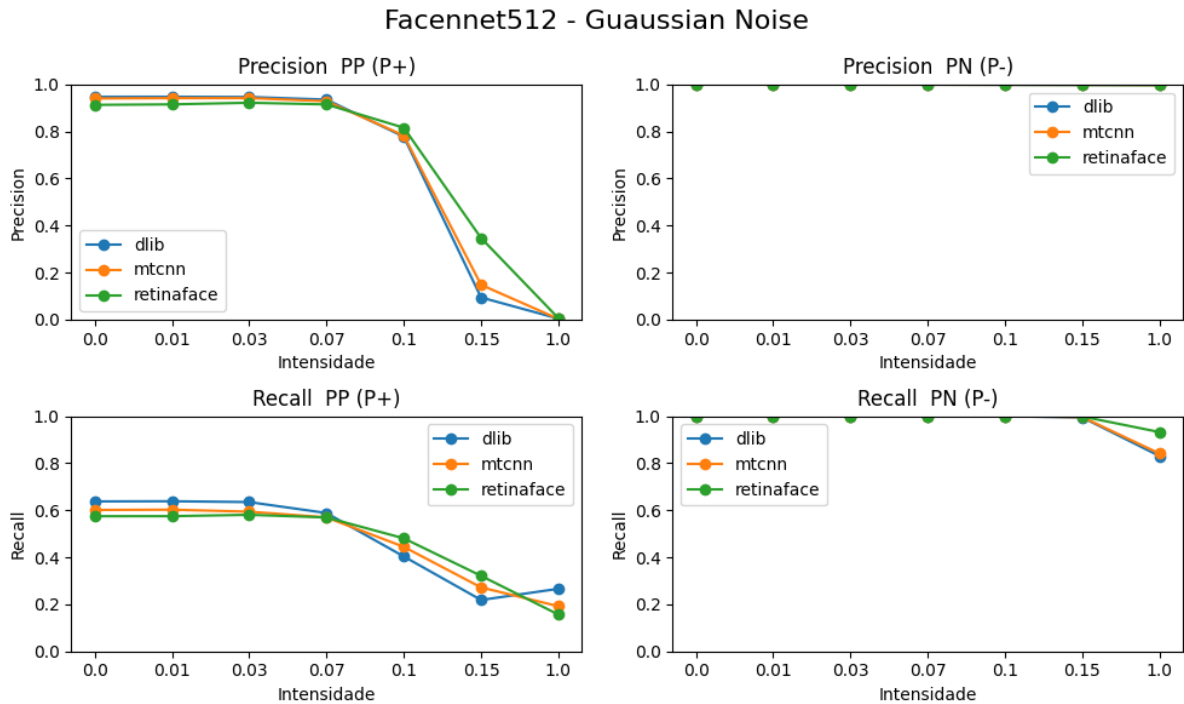


Figura I.4: Borramento Gaussiano - Facenet512

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo Facenet512 em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

### I.1.2 Ruído Gaussiano

No contexto da degradação por ruído gaussiano, o algoritmo ArcFace demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do ruído aumenta.

No contexto da degradação por ruído gaussiano, o algoritmo ArcFace demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do ruído aumenta.

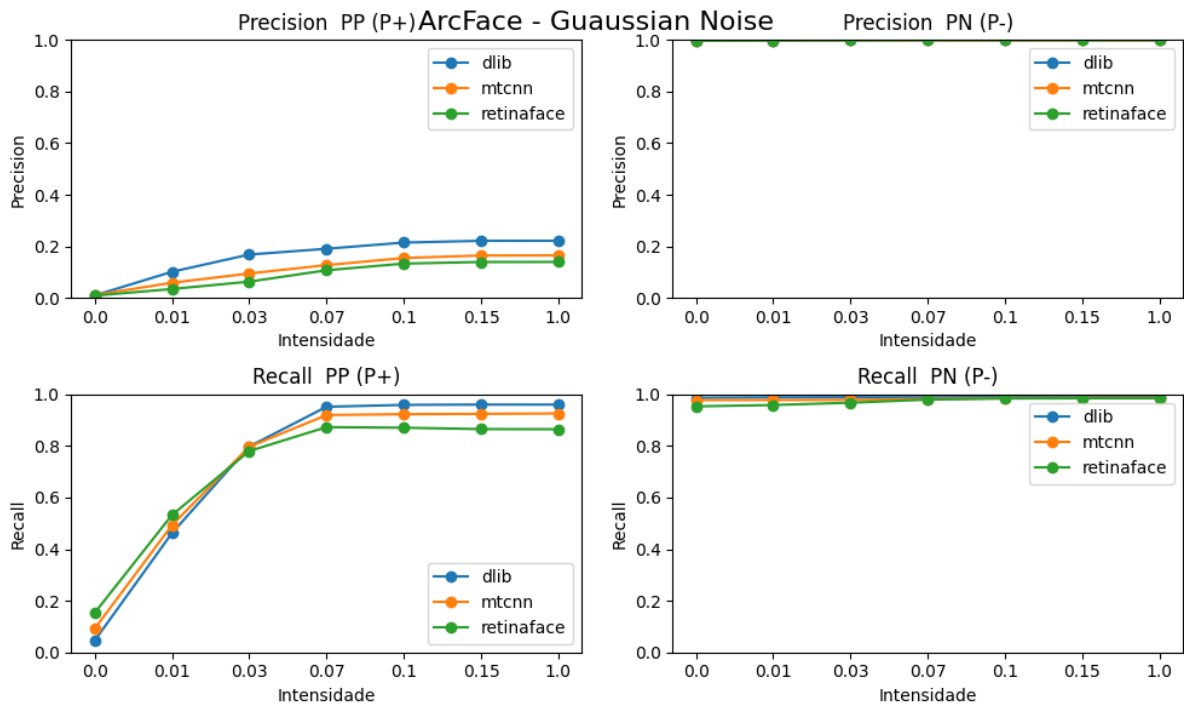


Figura I.5: Ruído Gaussiano - ArcFace

### I.1.3 Redimensionado

A degradação por redimensionamento demonstrou baixo desempenho em comparação aos demais algoritmos de reconhecimento facial para o modelo ArcFace.

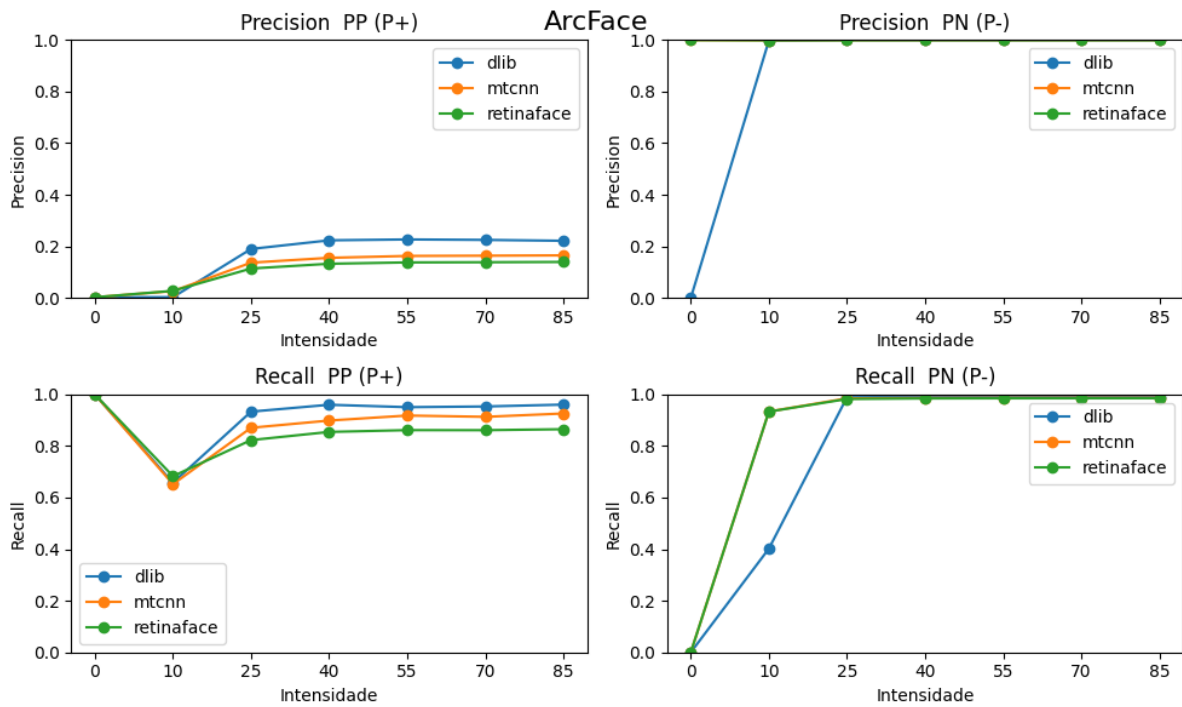


Figura I.6: Redimensionado - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original. No contexto da degradação por redimensionamento, o algoritmo ArcFace demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do redimensionamento aumenta.

#### I.1.4 Escurecimento

A degradação por escurecimento demonstrou baixa acurácia para os seguintes modelos:

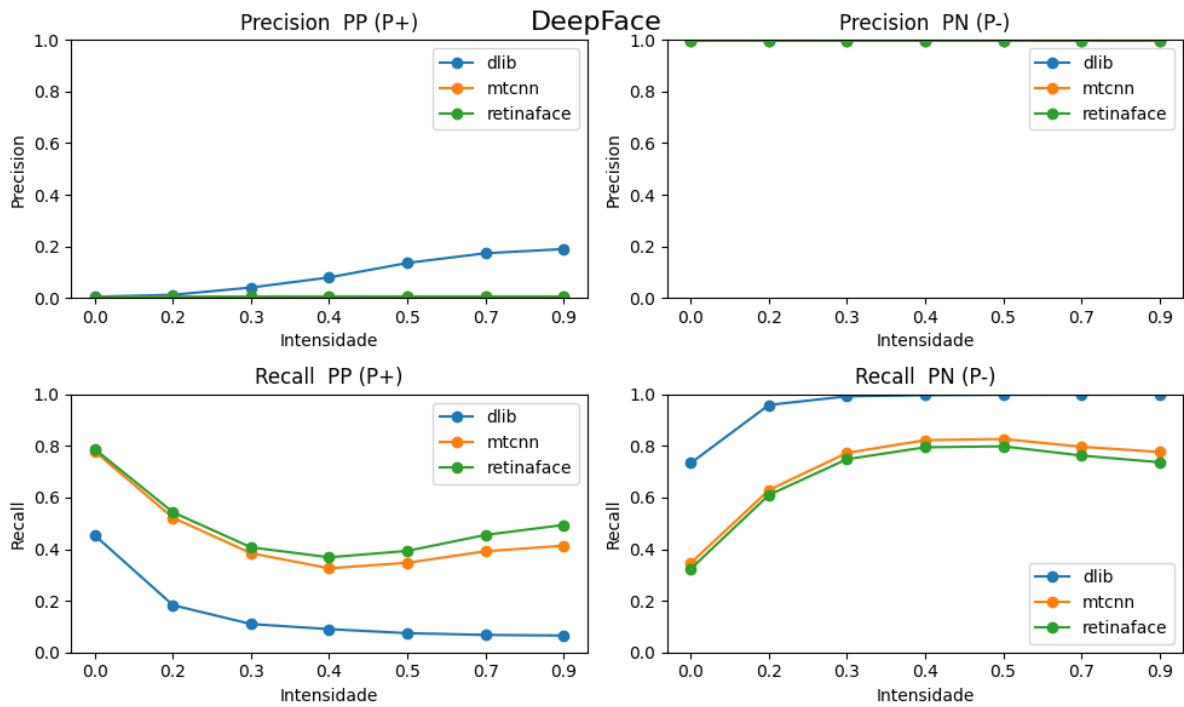


Figura I.7: Escurecimento - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original. No contexto da degradação por escurecimento, o algoritmo DeepFace demonstrou baixa eficácia em lidar com essa degradação. Isso pode ser observado nas curvas de precisão e taxa de recuperação, que apresentam uma queda acentuada à medida que a intensidade do escurecimento aumenta.

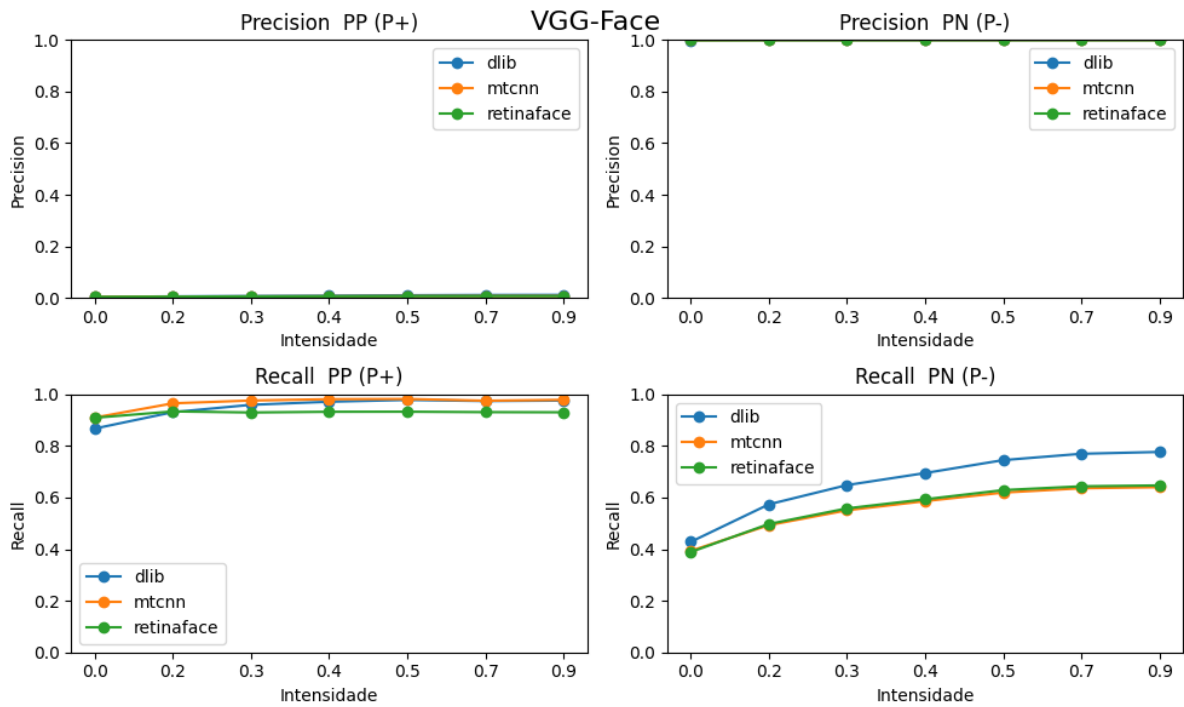


Figura I.8: Escurecimento - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

### I.1.5 Desfoque de Movimento

Nesta degradação, o algoritmo Arface e VGG-Face demonstraram baixo desempenho em comparação aos demais algoritmos de reconhecimento facial.

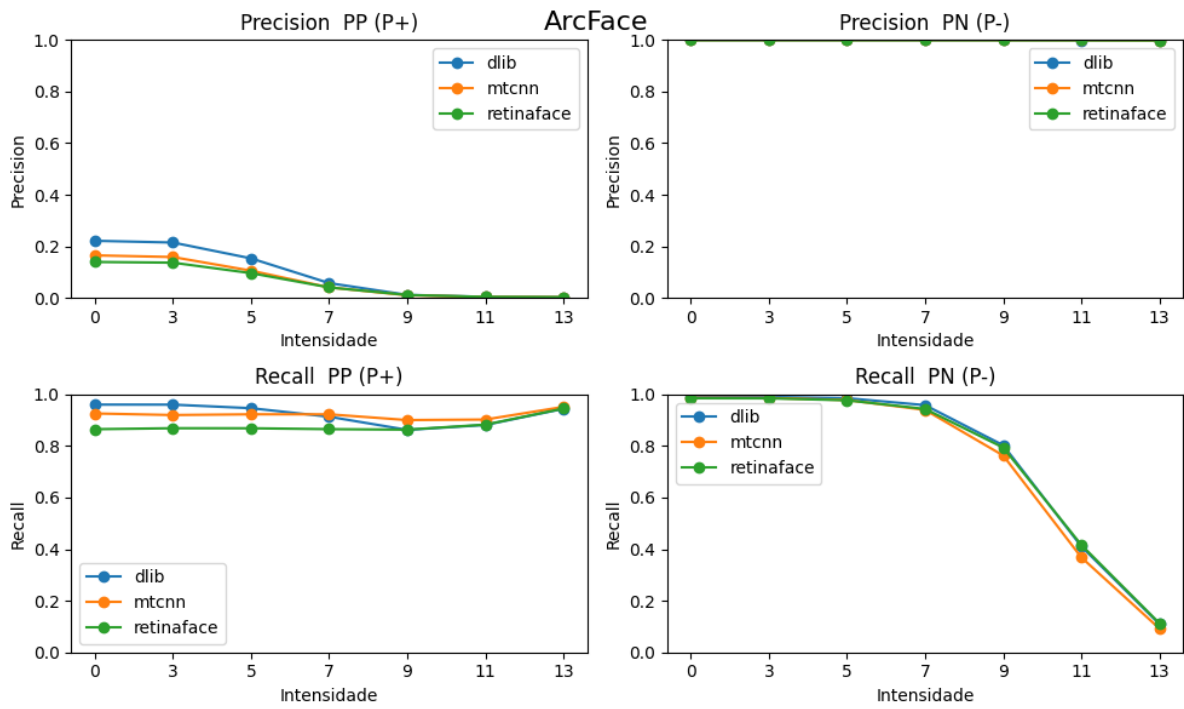


Figura I.9: Desfoque de Movimento - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original



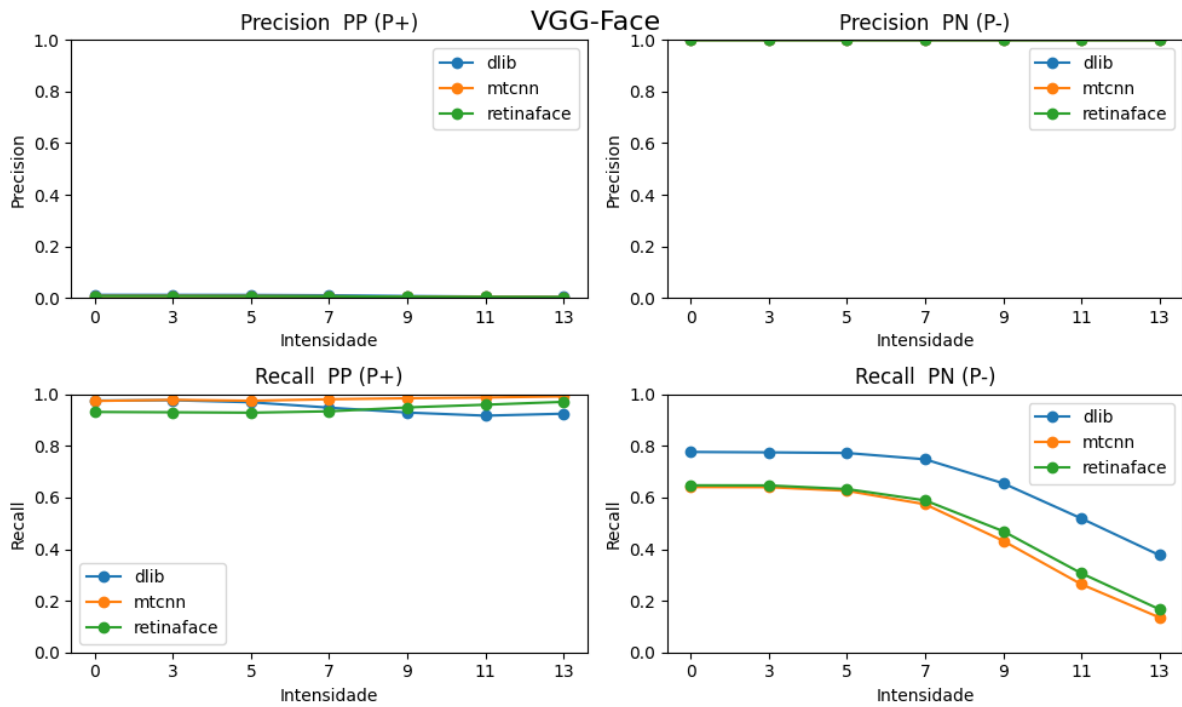


Figura I.10: Desfoque de Movimento - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

### I.1.6 Ruído de Sal e Pimenta

Nesta degradação, o algoritmo Arface e VGG-Face demonstraram baixo desempenho em comparação aos demais algoritmos de reconhecimento facial.

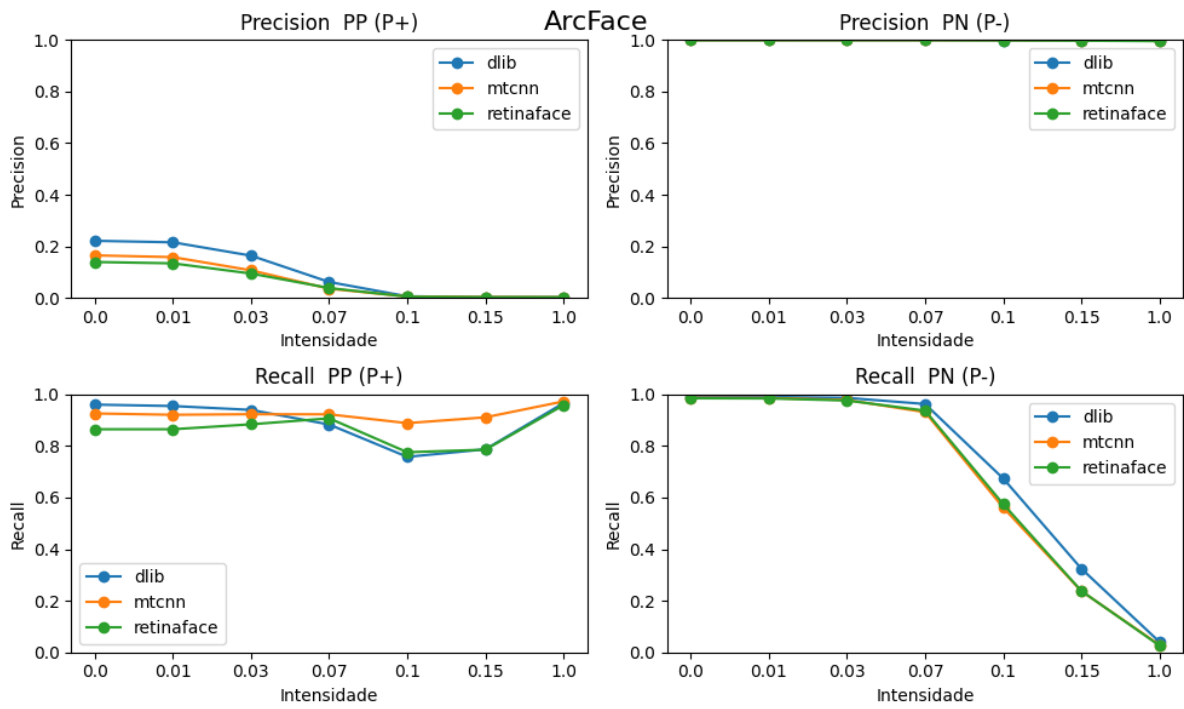


Figura I.11: Ruído de Sal e Pimenta - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

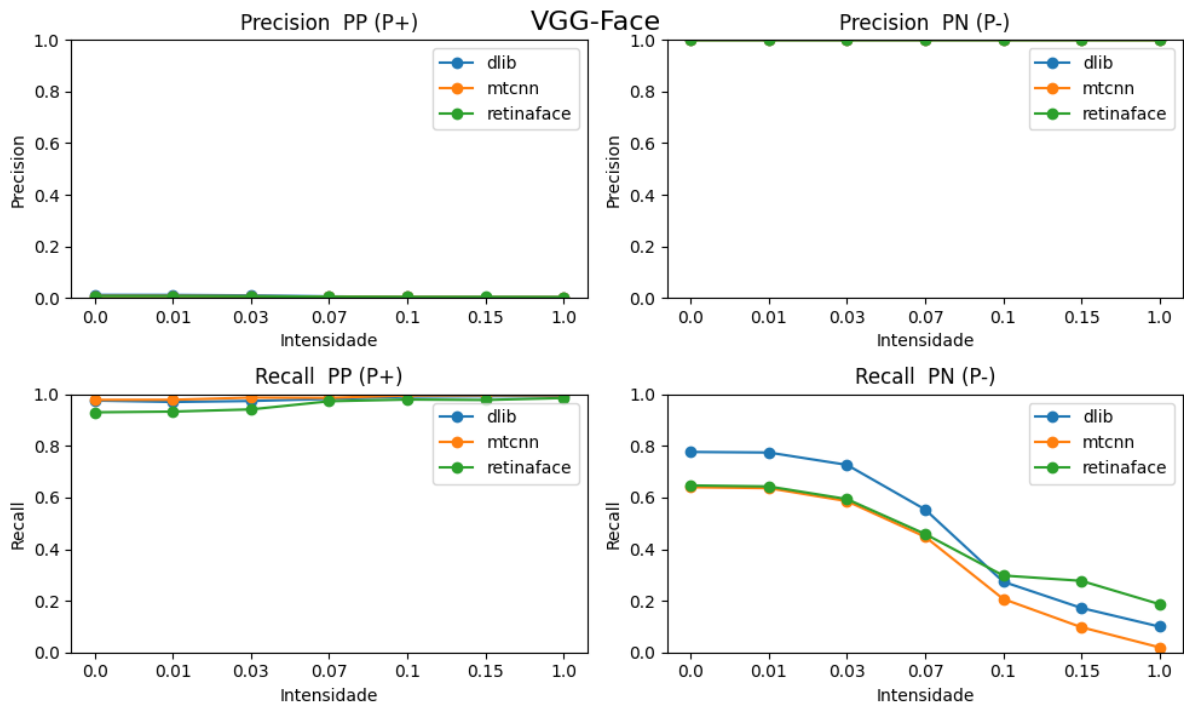


Figura I.12: Ruído de Sal e Pimenta - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

### I.1.7 Compressão JPEG

Nesta degradação, o algoritmo VGG-Face demonstrou baixo desempenho em comparação aos demais algoritmos de reconhecimento facial.

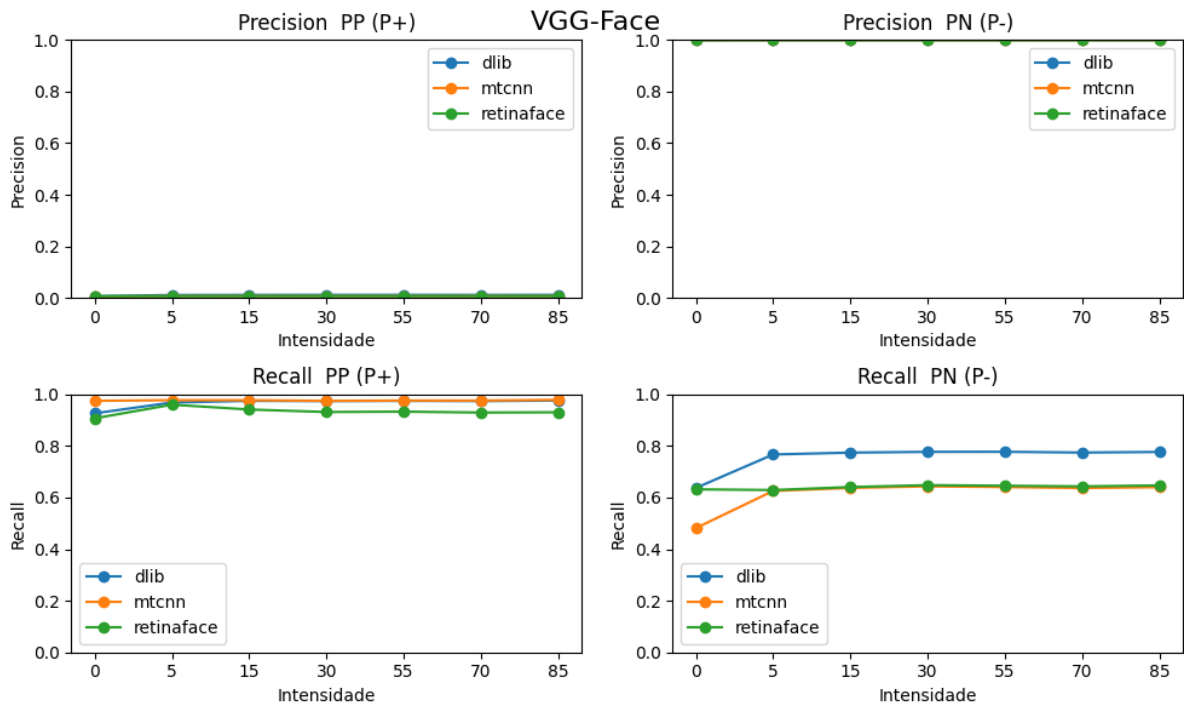


Figura I.13: Compressão JPEG - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

## I.2 No contexto do Teste 2

### Imagem original x degradada

Nas métricas geradas pelo teste 2, foi possível observar que os algoritmos ArcFace e VGG-Face demonstraram baixo desempenho em relação aos demais algoritmos de reconhecimento facial. Isso pode ser observado nos gráficos abaixo.

#### I.2.1 Ruído Gaussiano

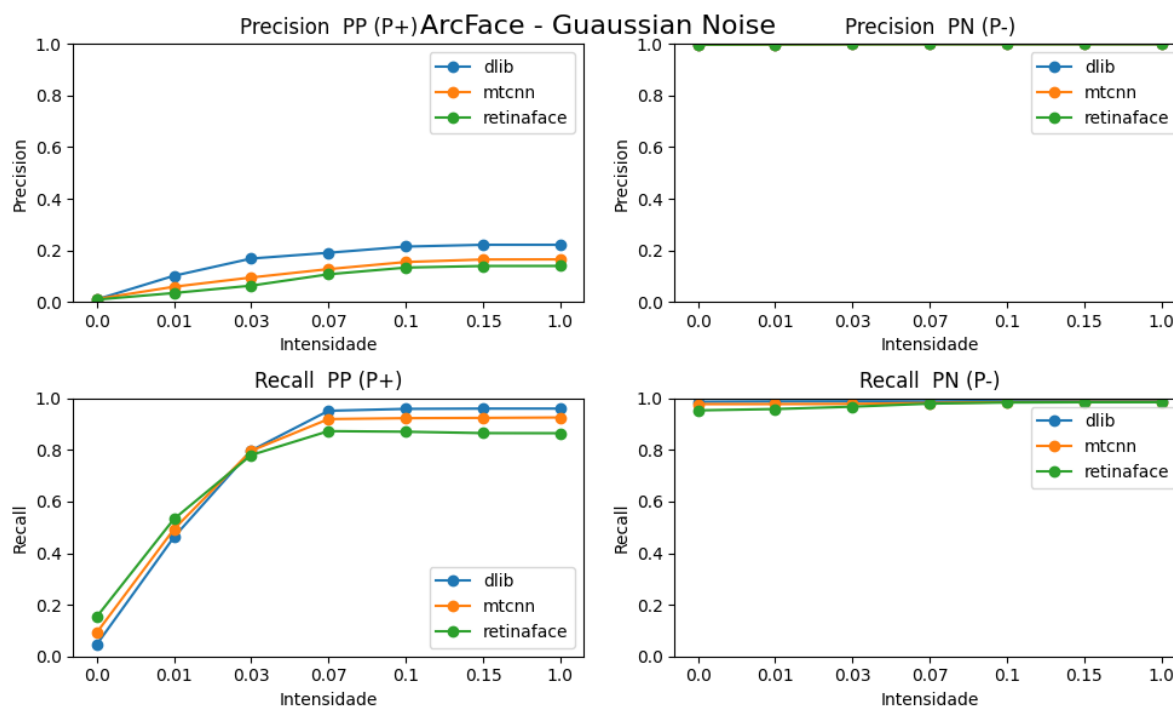


Figura I.14: Ruído Gaussiano - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

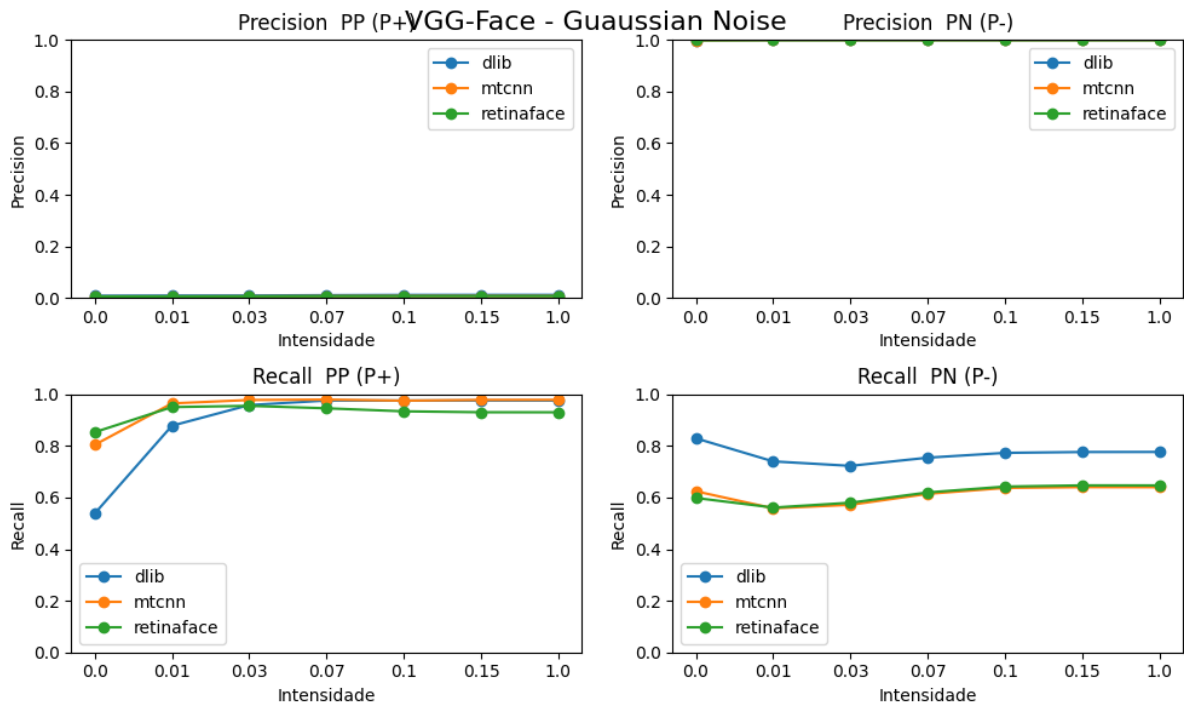


Figura I.15: Ruído Gaussiano - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

## I.2.2 Borramento Gaussiano

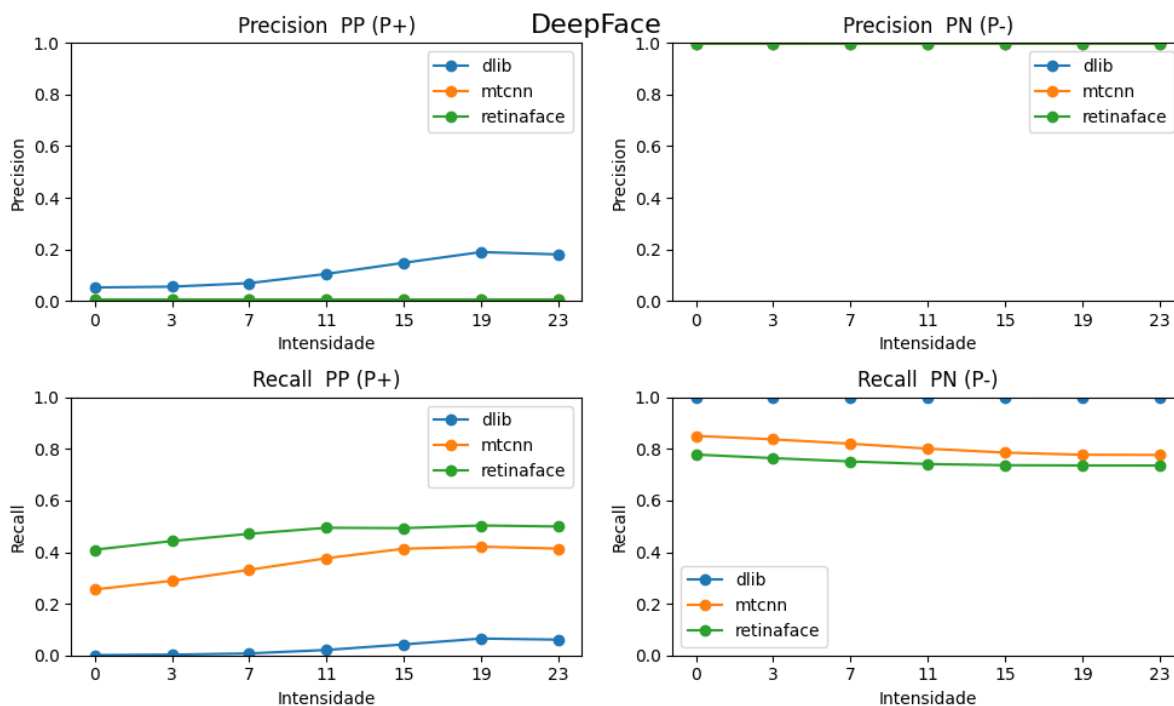


Figura I.16: Borramento Gaussiano - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

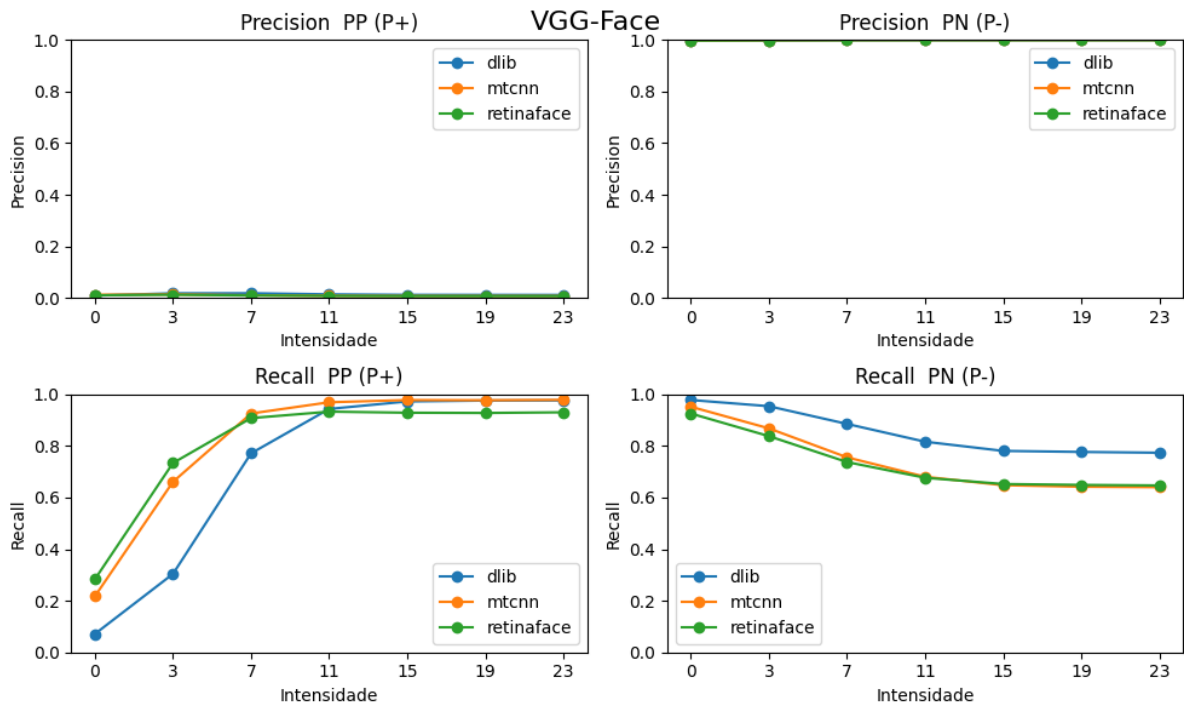


Figura I.17: Borramento Gaussiano - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original



### I.2.3 Escurecimento

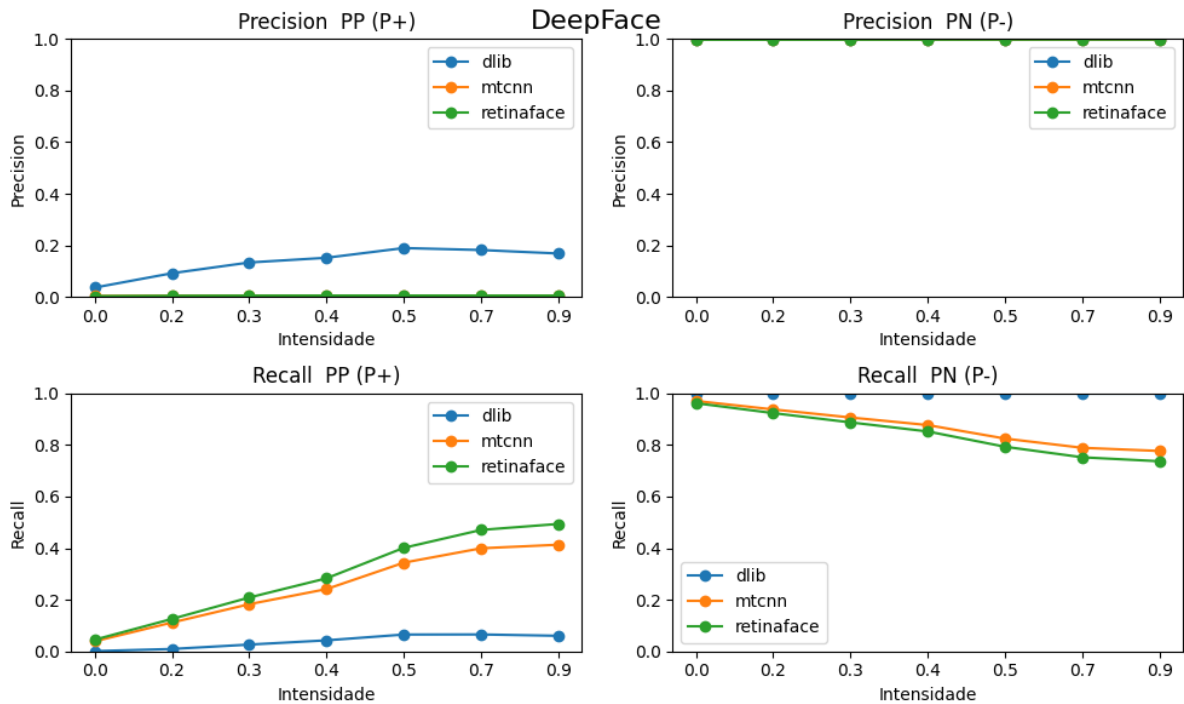


Figura I.18: Escurecimento - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

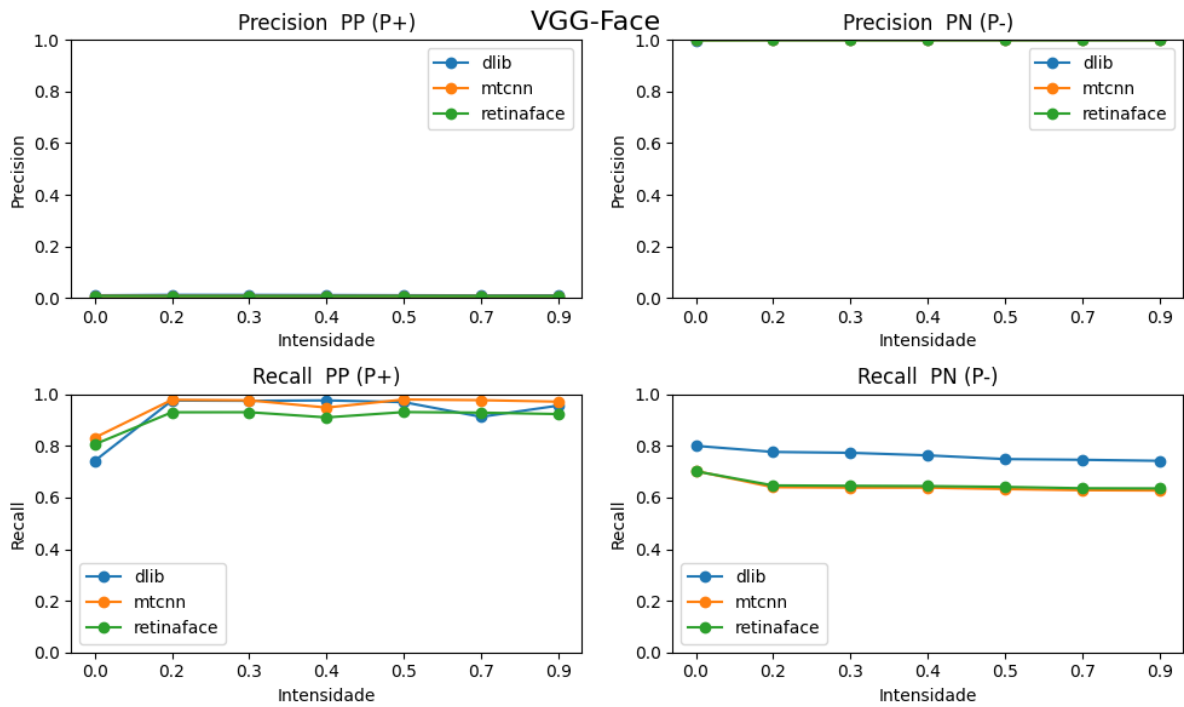


Figura I.19: Escurecimento - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

## I.2.4 Desfoque de Movimento

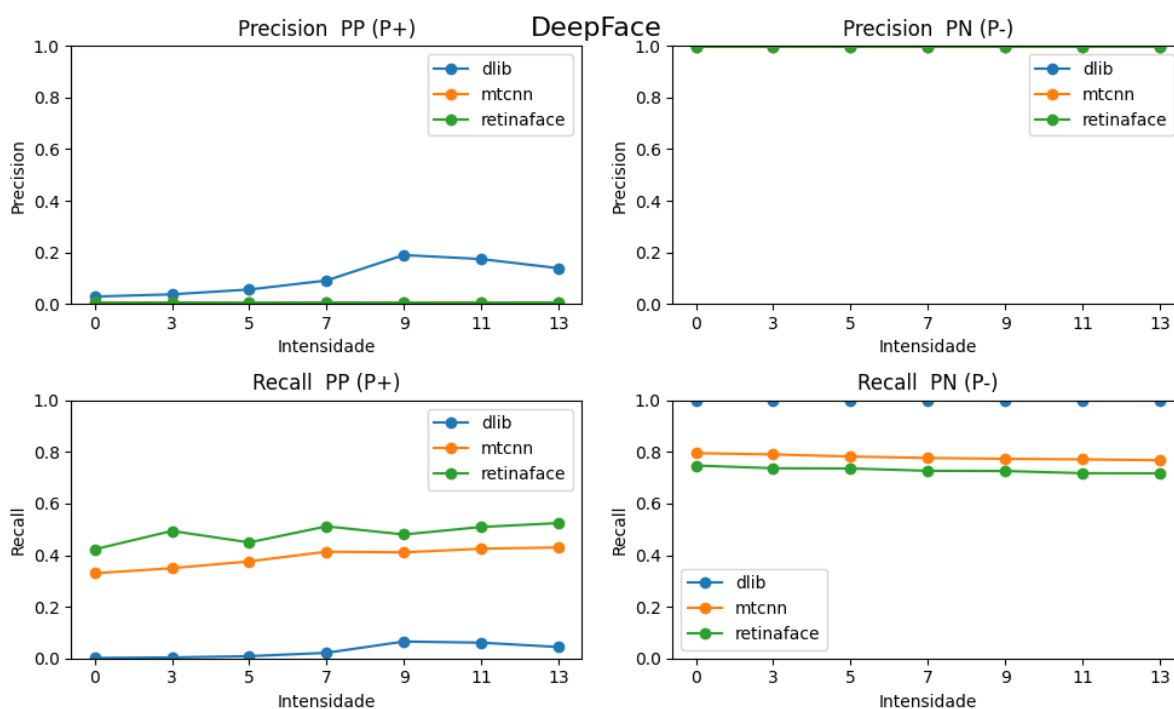


Figura I.20: Desfoque de Movimento - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

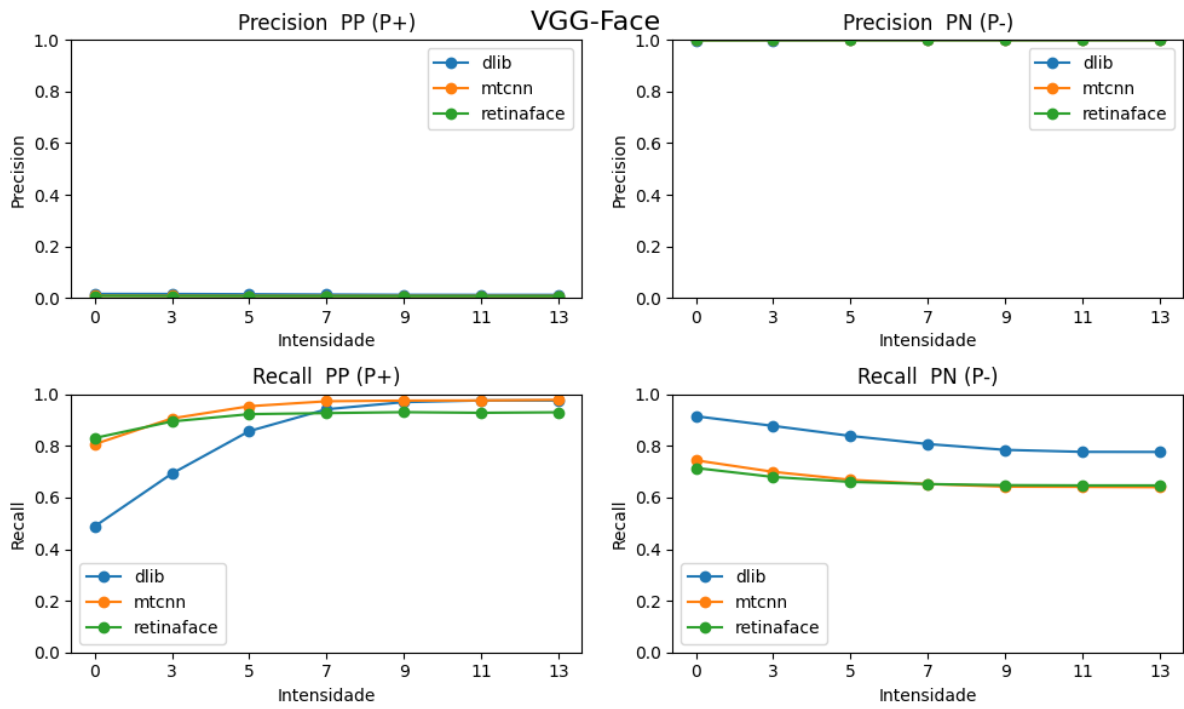


Figura I.21: Desfoque de Movimento - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

## I.2.5 Ruído de Sal e Pimenta

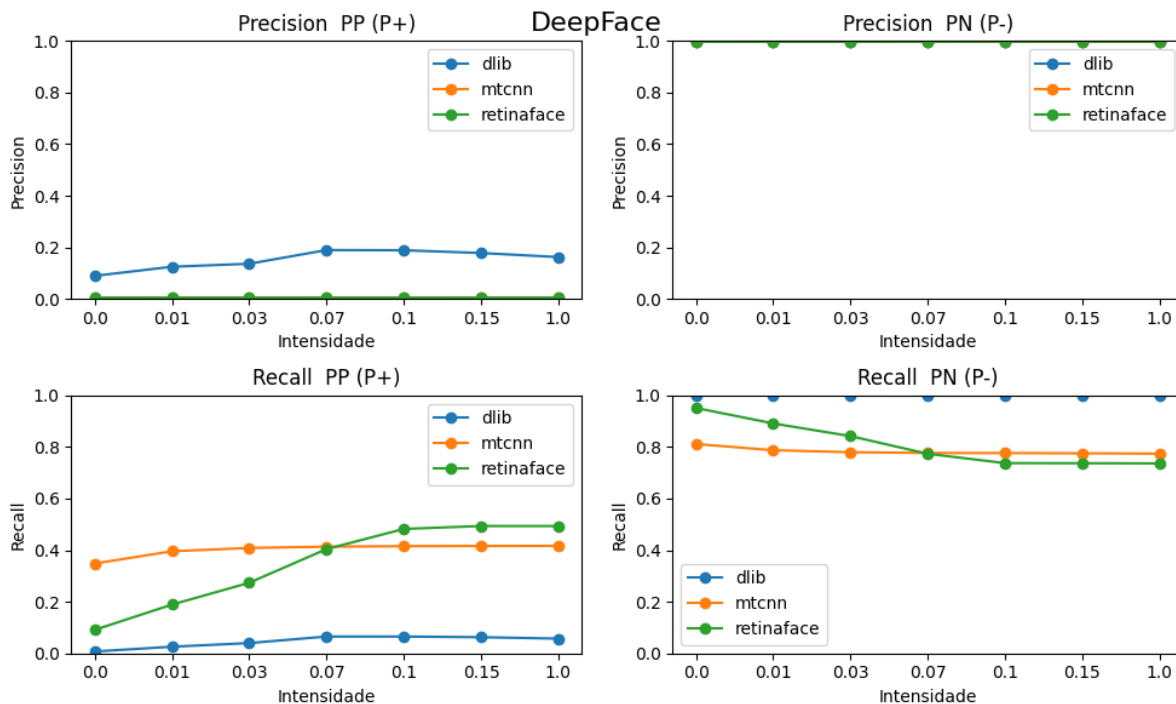


Figura I.22: Ruído de Sal e Pimenta - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

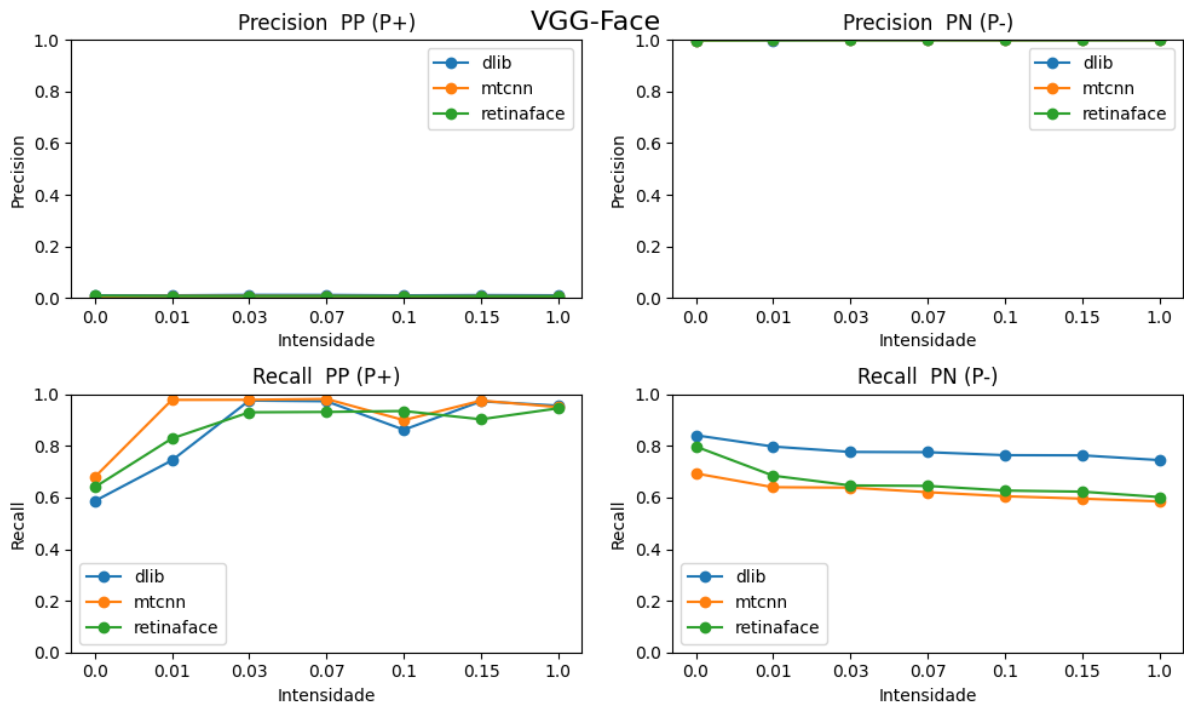


Figura I.23: Ruído de Sal e Pimenta - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

## I.2.6 Compressão JPEG

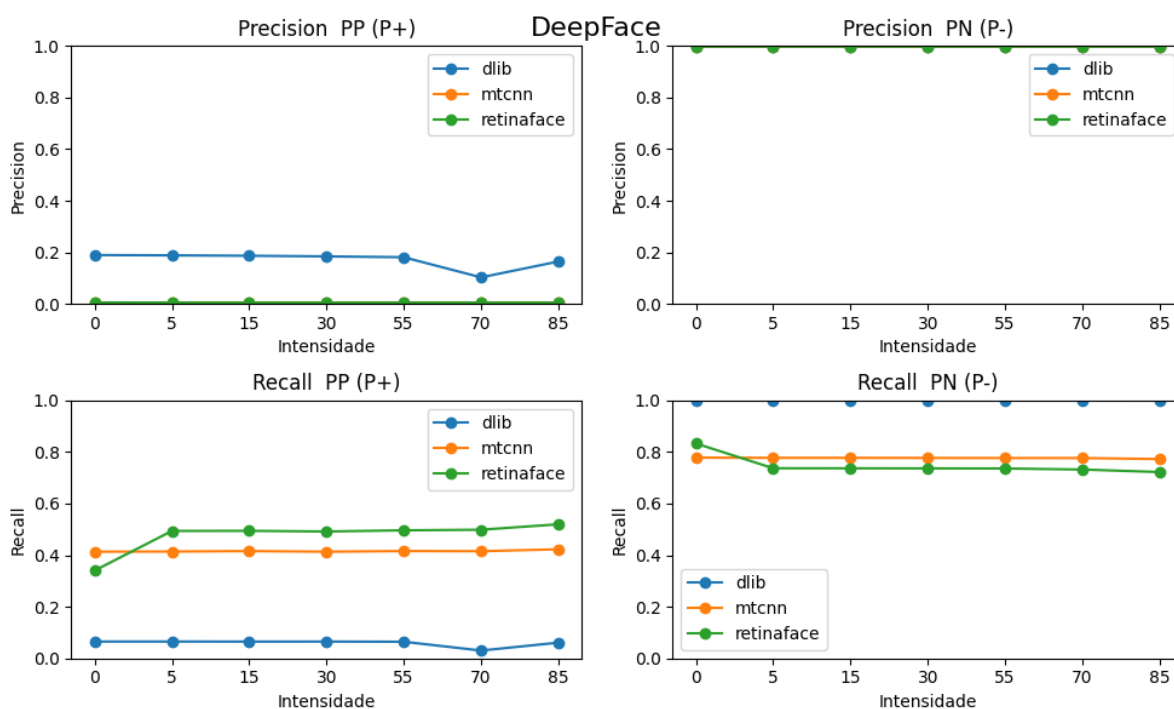


Figura I.24: Compressão JPEG - DeepFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepFace em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

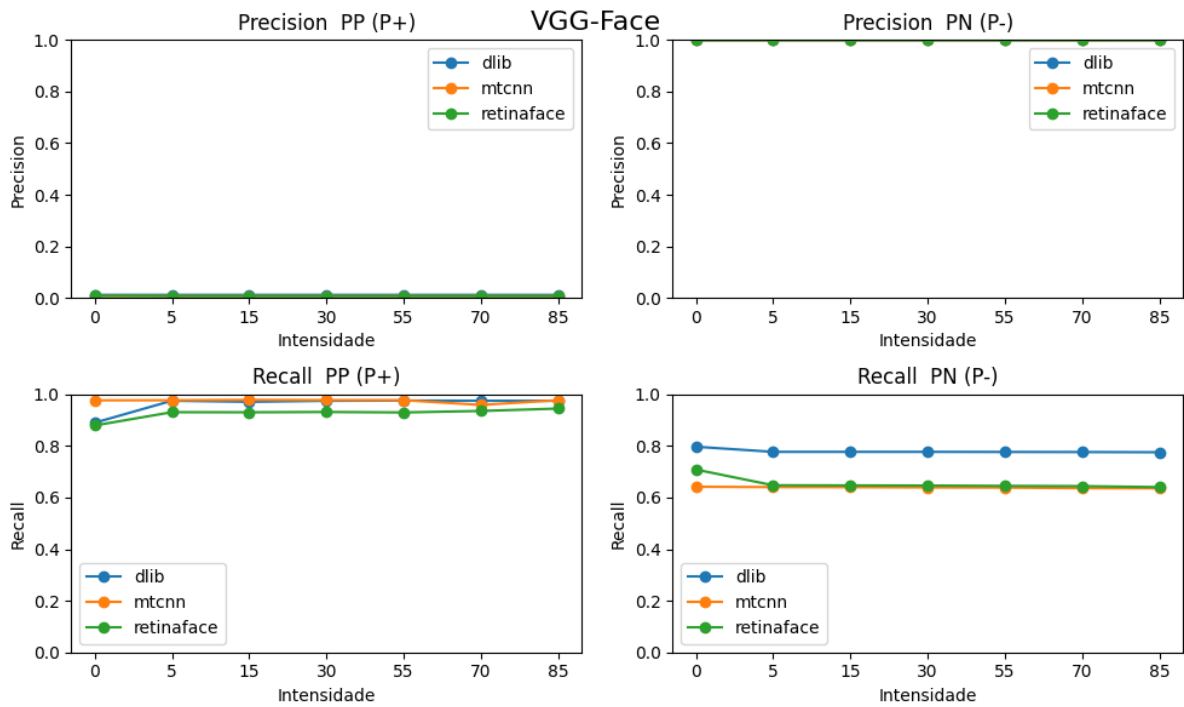


Figura I.25: Compressão JPEG - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem degradada até a intensidade máxima. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original



## I.3 No contexto do Teste 3

### Imagem original x mesma imagem original

Nessa combinação de imagens, a imagem original é comparada com ela mesma adicionada as intensidades de degradação. O objetivo é verificar se o algoritmo é capaz de identificar a mesma imagem com diferentes intensidades de degradação.

#### I.3.1 Ruído Gaussiano

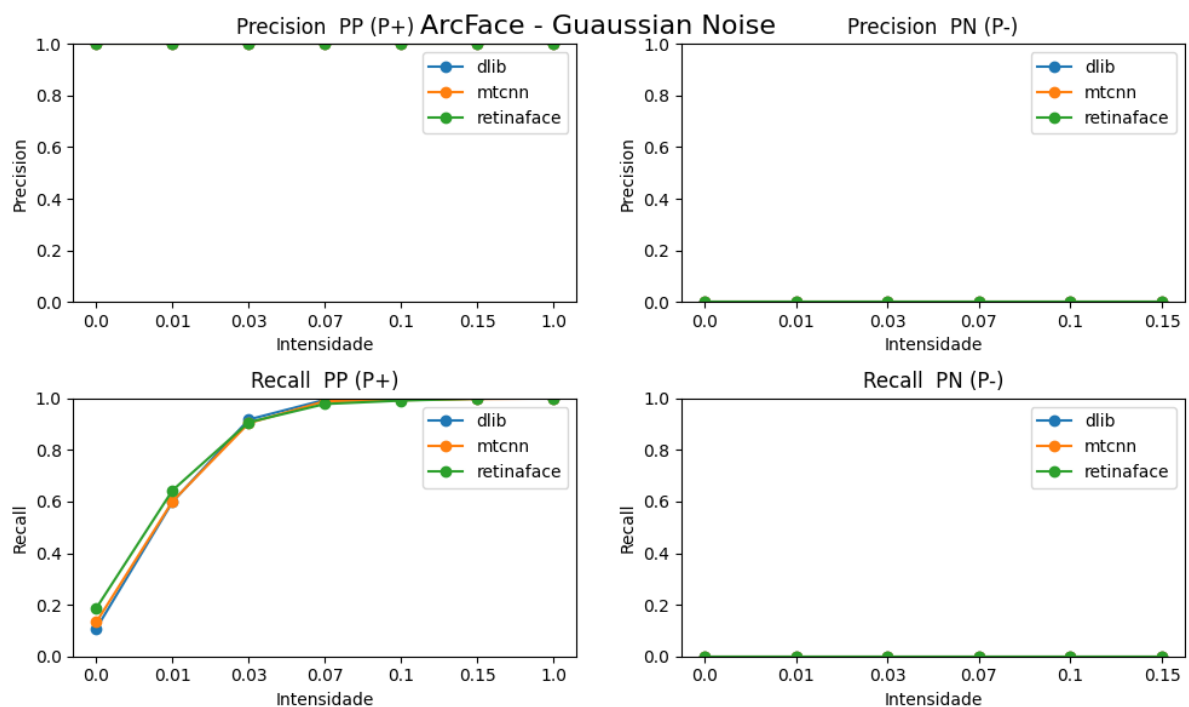


Figura I.26: Ruído Gaussiano - ArcFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo ArcFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

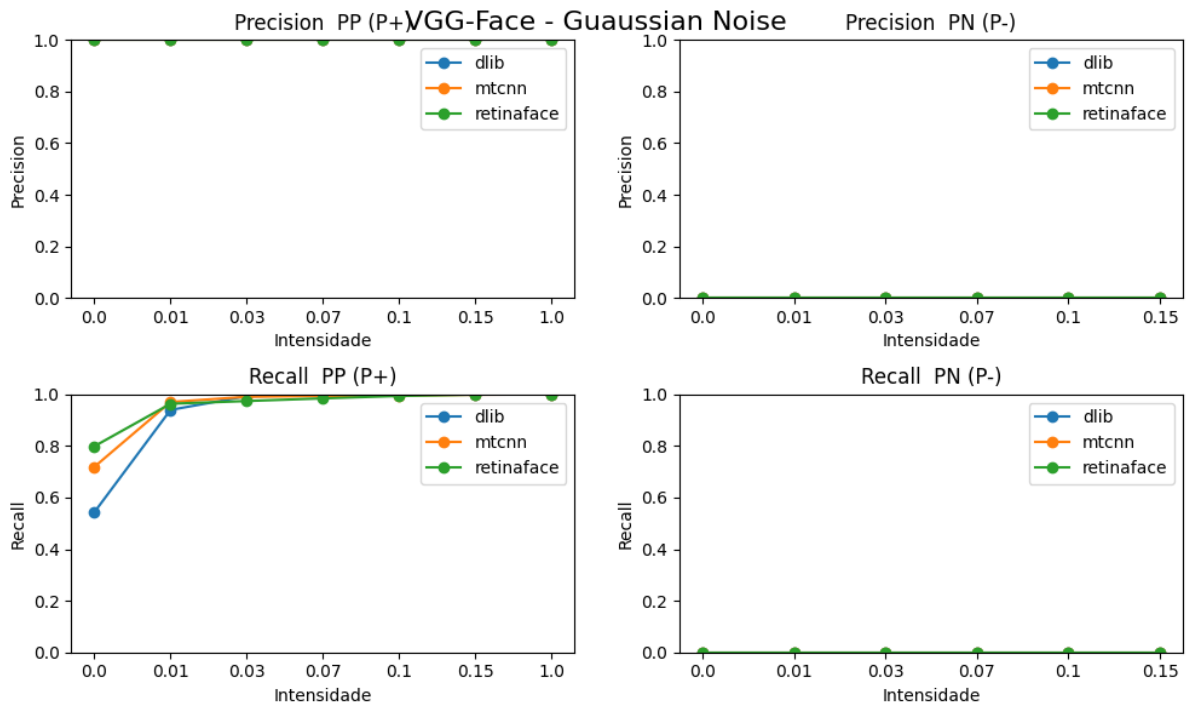


Figura I.27: Ruído Gaussiano - VGG-Face

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo VGG-Face em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

### I.3.2 Borramento Gaussiano

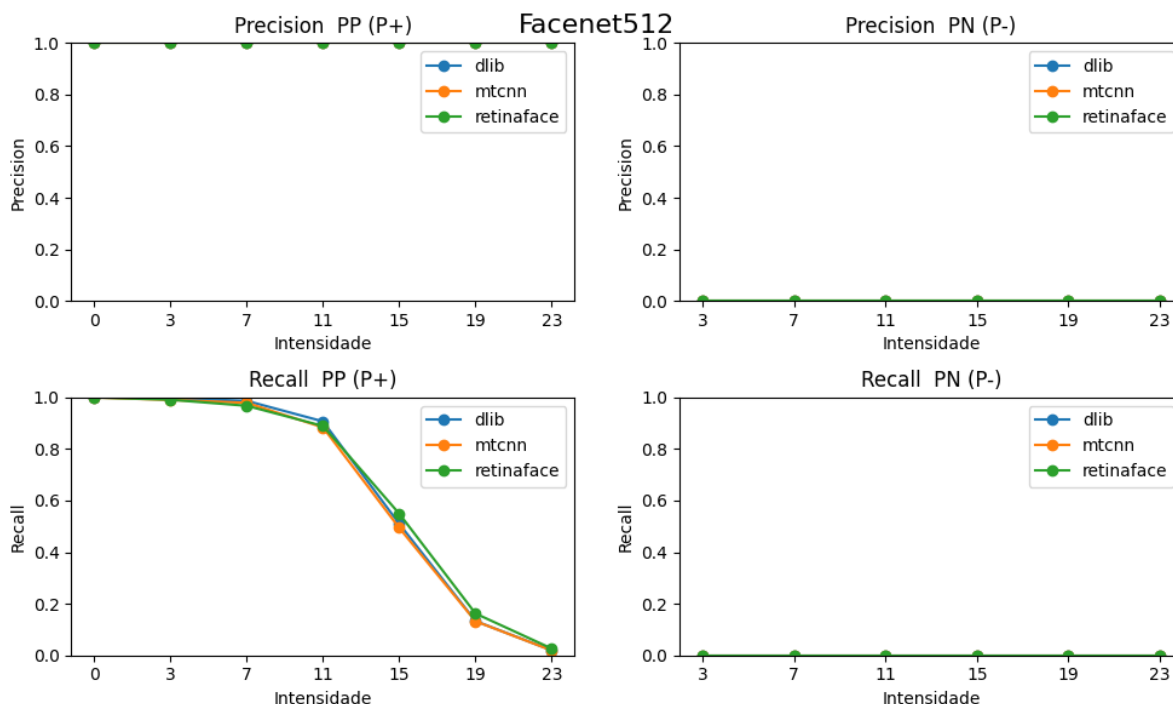


Figura I.28: Borramento Gaussiano - Facenet512

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo Facenet512 em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

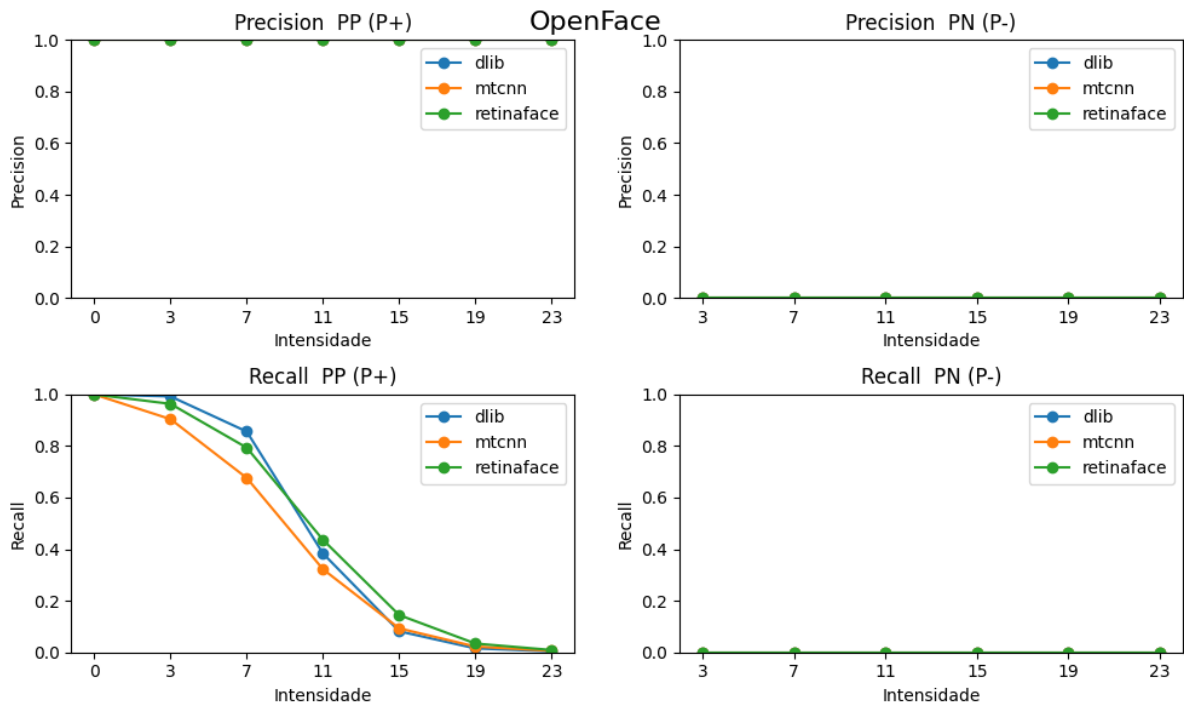


Figura I.29: Borramento Gaussiano - OpenFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo OpenFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original

### I.3.3 Escurecimento

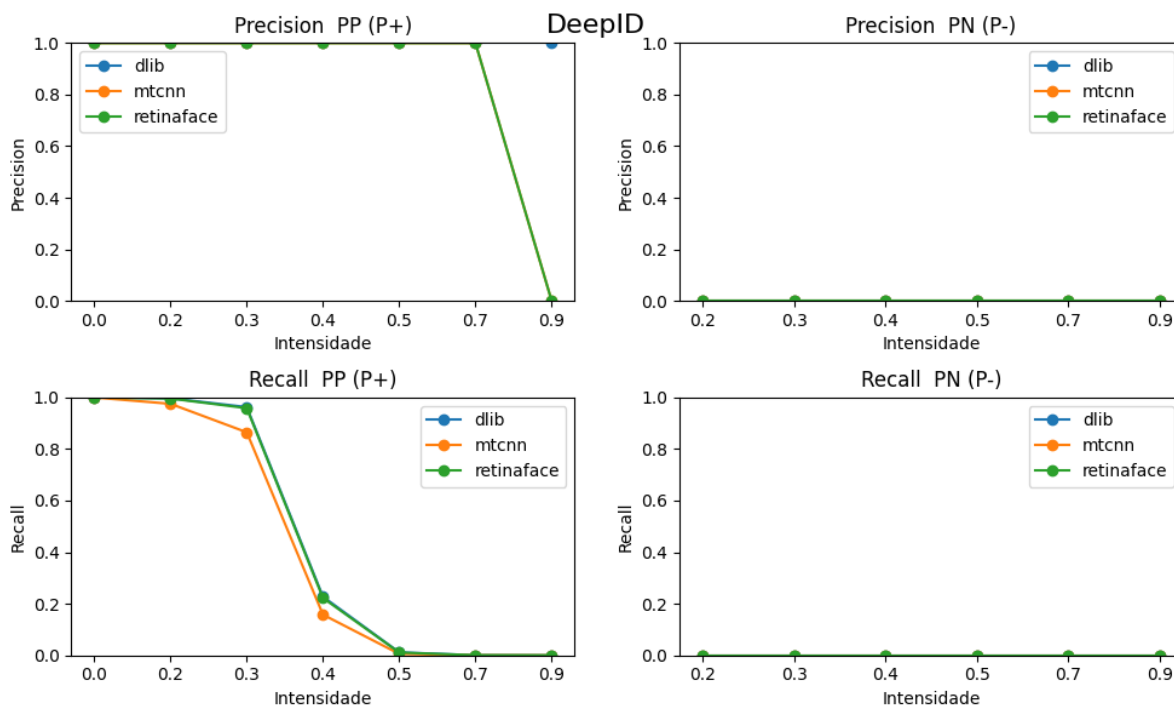


Figura I.30: Escurecimento - DeepID

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo DeepID em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

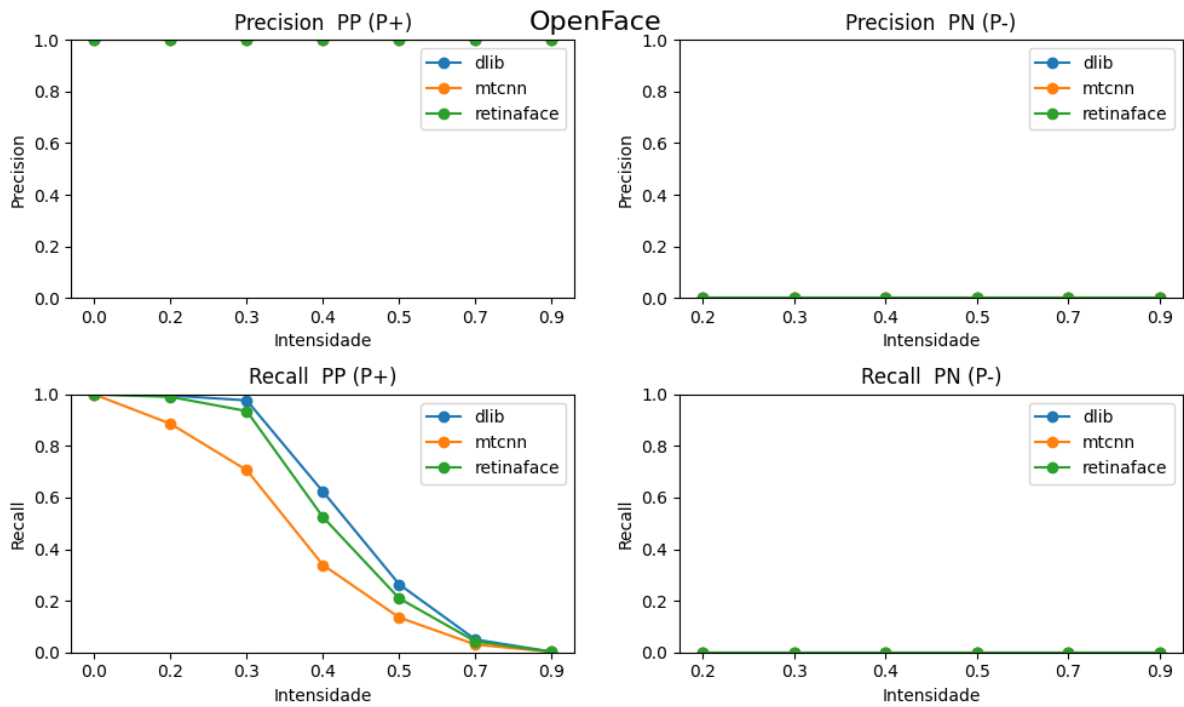


Figura I.31: Escurecimento - OpenFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo OpenFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

### I.3.4 Desfoque de Movimento

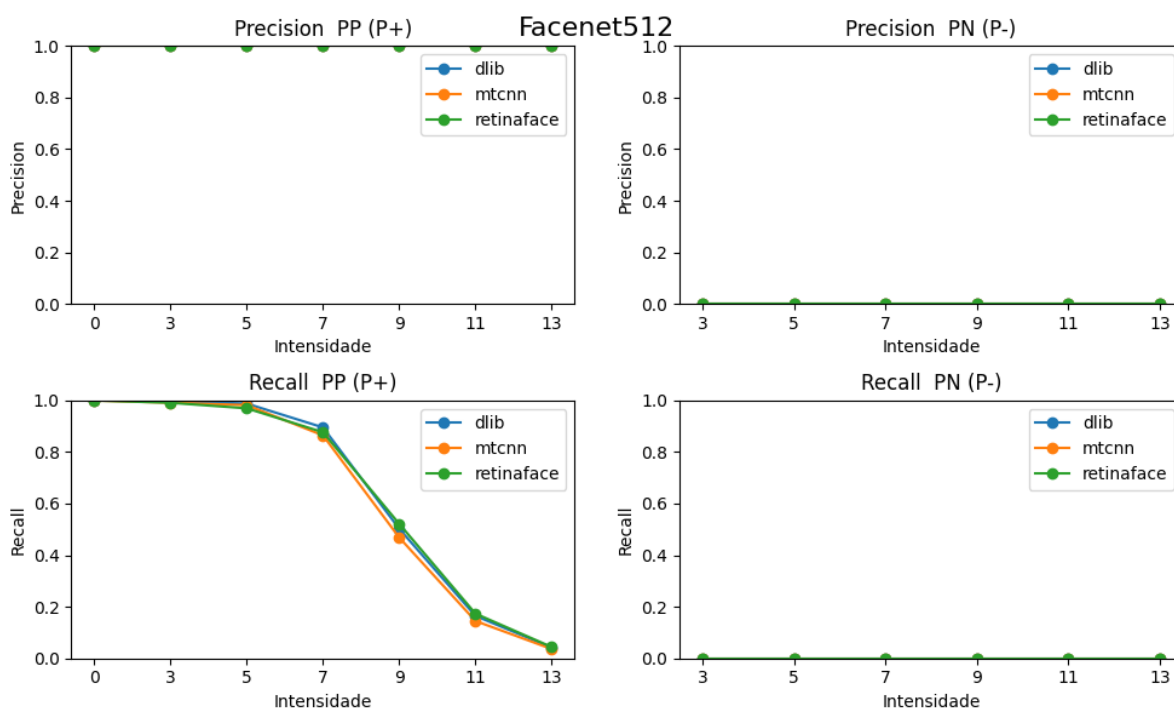


Figura I.32: Desfoque de Movimento - Facenet512

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo Facenet512 em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

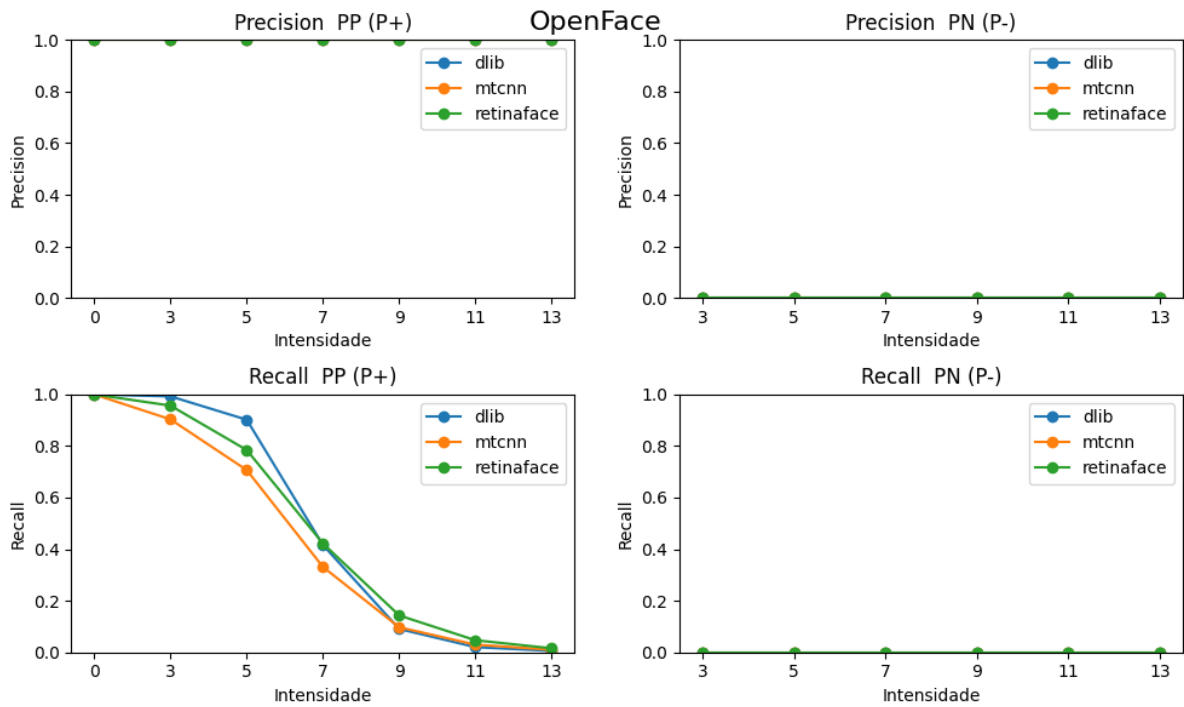


Figura I.33: Desfoque de Movimento - OpenFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo OpenFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.



### I.3.5 Ruído de Sal e Pimenta

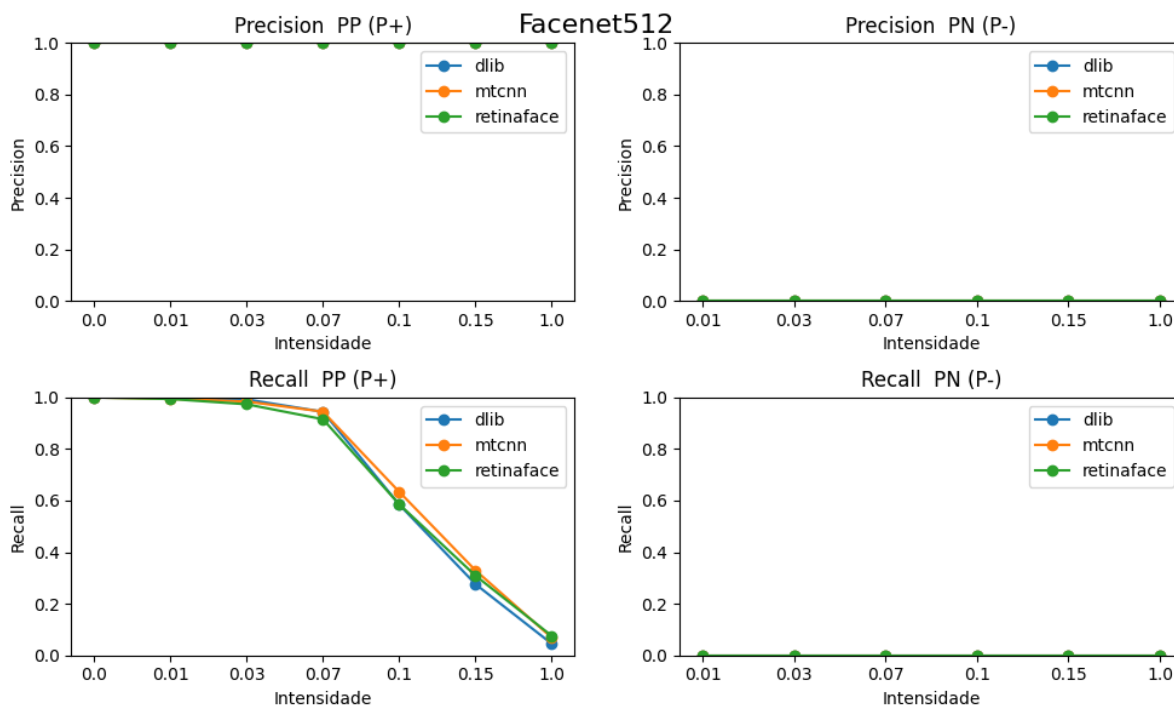


Figura I.34: Ruído de Sal e Pimenta - Facenet512

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo Facenet512 em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

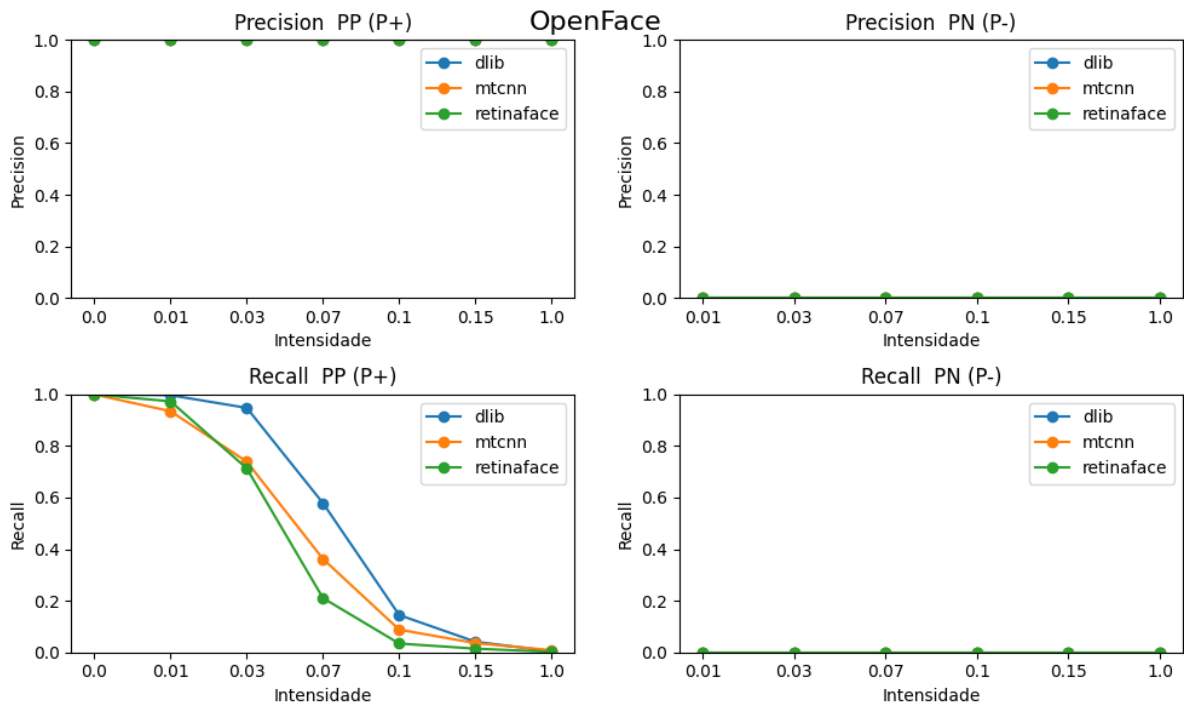


Figura I.35: Ruído de Sal e Pimenta - OpenFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo OpenFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

### I.3.6 Compressão JPEG

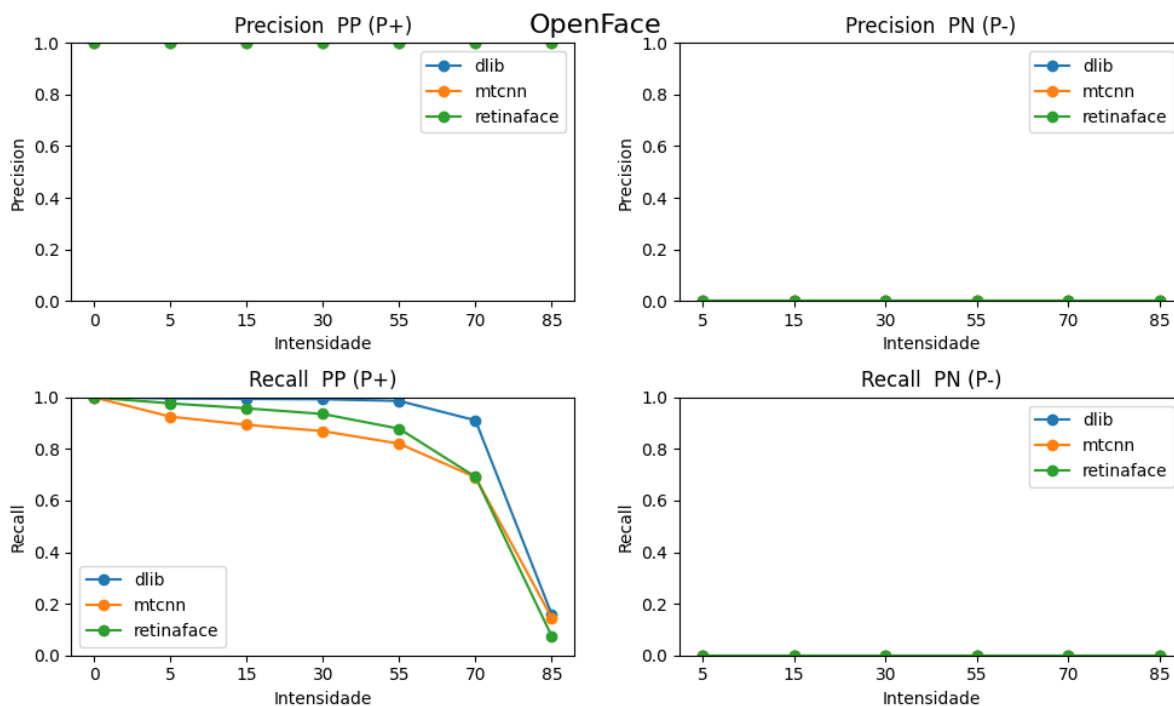


Figura I.36: Compressão JPEG - OpenFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo OpenFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.

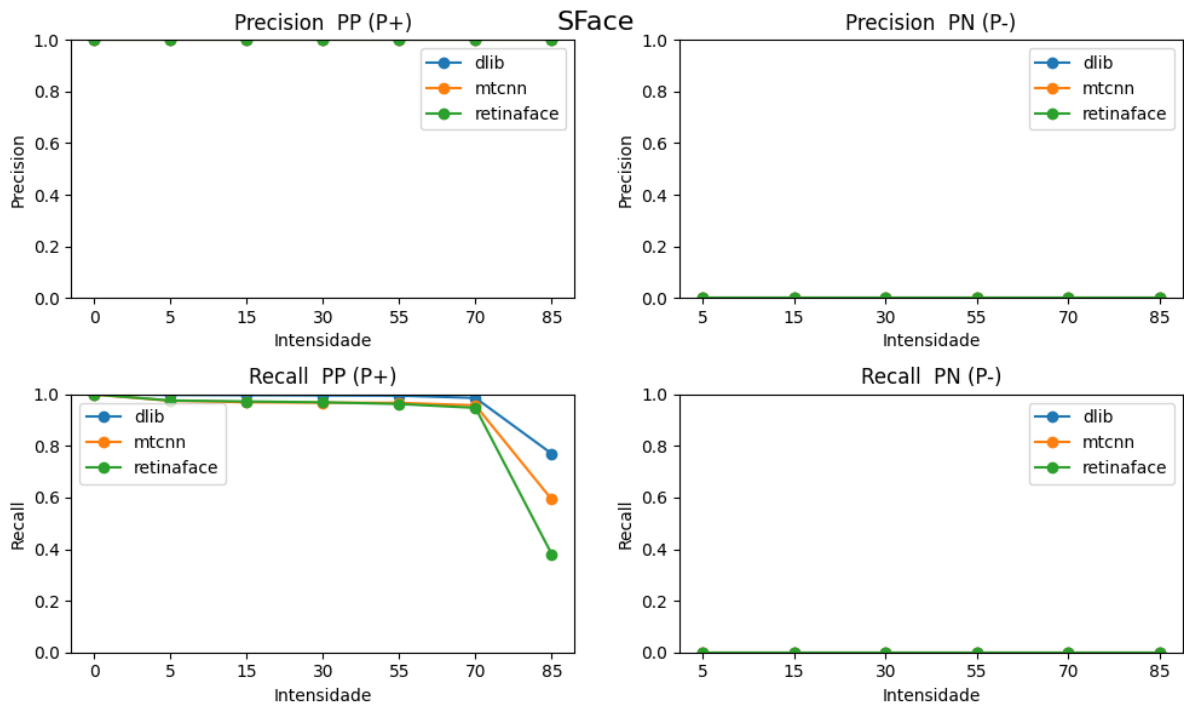


Figura I.37: Compressão JPEG - SFace

O gráfico apresenta a comparação dos resultados obtidos pelo algoritmo SFace em relação à imagem degradada e à imagem original. A imagem original é comparada com ela mesma adicionada as intensidades de degradação. A intensidade máxima representa o nível máximo de degradação aplicado à imagem original.