



Universidade Federal de Roraima  
Departamento de Ciência da Computação  
Bacharelado em Ciência da Computação



# GAN-AGE: Um sistema para geração de imagens baseado em previsão temporal para pessoas desaparecidas

---

João Paulo Parreira Peixoto

Orientador: Prof. Dr. Herbert Oliveira Rocha

Boa Vista – RR

22 de Outubro de 2021

- 1. Introdução**
2. Fundamentação Teórica
3. Trabalhos Correlatos
4. Solução Proposta
5. Avaliação Experimental
6. Considerações Finais e Trabalhos Futuros

## Geração de Personagens de animes



Fonte: (JIN et al., 2017)

## Dando Vida a Criações Artísticas



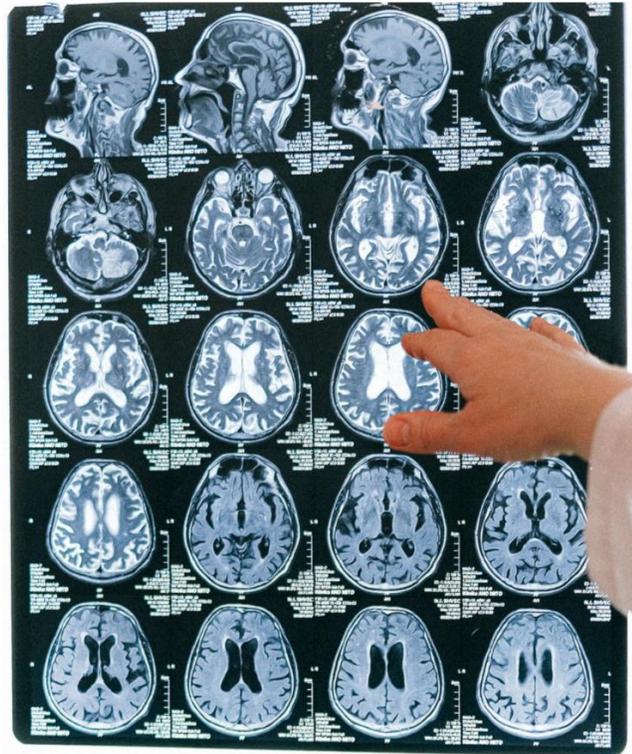
Fonte: (ZAKHAROV et al., 2019)

Aplicação baseada em tecnologias *Deep Learning* desenvolvida pela IBM

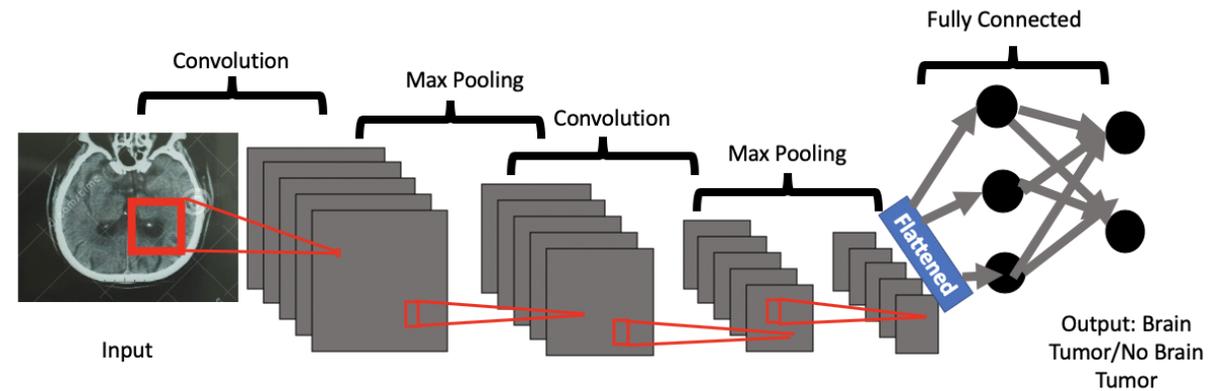


Fonte: (IBM, 2021)

## Detecção de doenças



## Rede CNN da aplicação



Fonte: (PIERRE, 2020)<sup>1</sup>

Fonte: (PIERRE, 2020)<sup>1</sup>

<sup>1</sup> <https://towardsdatascience.com/how-gans-can-improve-healthcare-analytics-7d2379eff19e>

- Crescimento do número de desaparecidos com uma taxa superior a 10% anualmente (VIEIRA, 2018).
- No Brasil, são registrados em média 50 mil desaparecimentos de crianças e adolescentes por ano (VIEIRA, 2018).
- Estimativa de 25 milhões de desaparecimentos no mundo (VIEIRA, 2018).
- Estima-se 460 mil crianças desaparecidas todos os anos nos Estados Unidos (GLOBALMISSINGKIDS, 2020).

O problema considerado neste trabalho pode ser expresso com a seguinte questão:

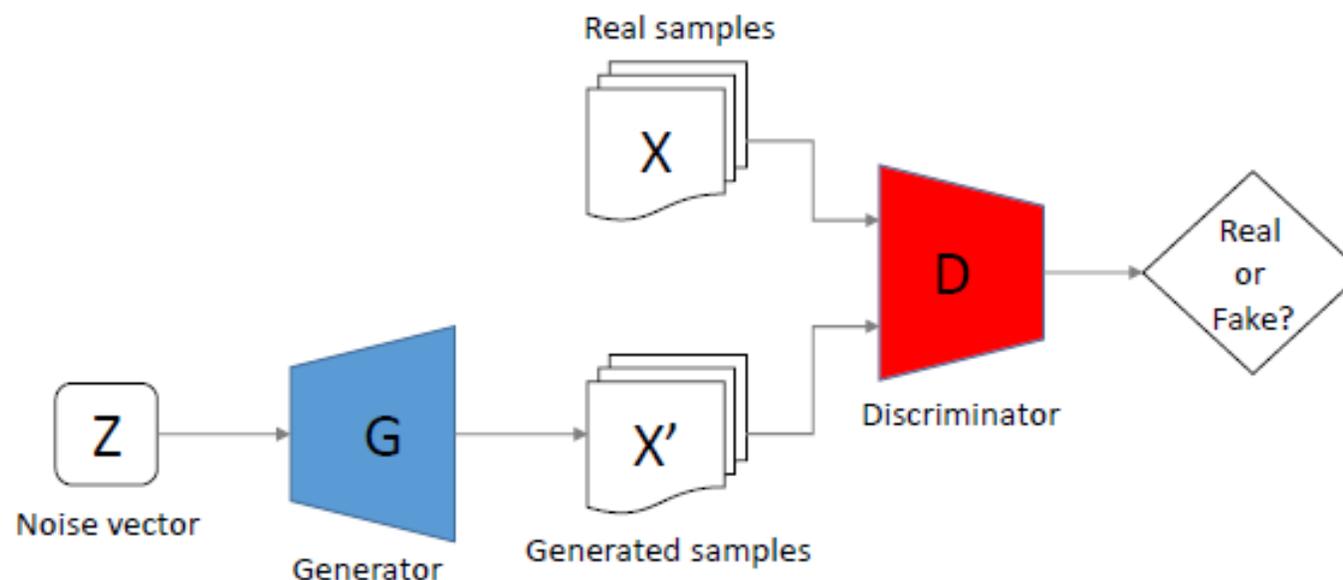
**Como uma imagem de um rosto humano pode ser envelhecida, considerando intervalos de idades por meio da utilização de GANs, para auxiliar em investigações de desaparecidos?**

Projetar, desenvolver e avaliar um modelo usando **GANs** para geração de imagens do **rosto humano envelhecido** baseado em intervalos de idades para uma estimativa na **previsão temporal** de pessoas desaparecidas, visando assim contribuir em investigações de **pessoas desaparecidas**.

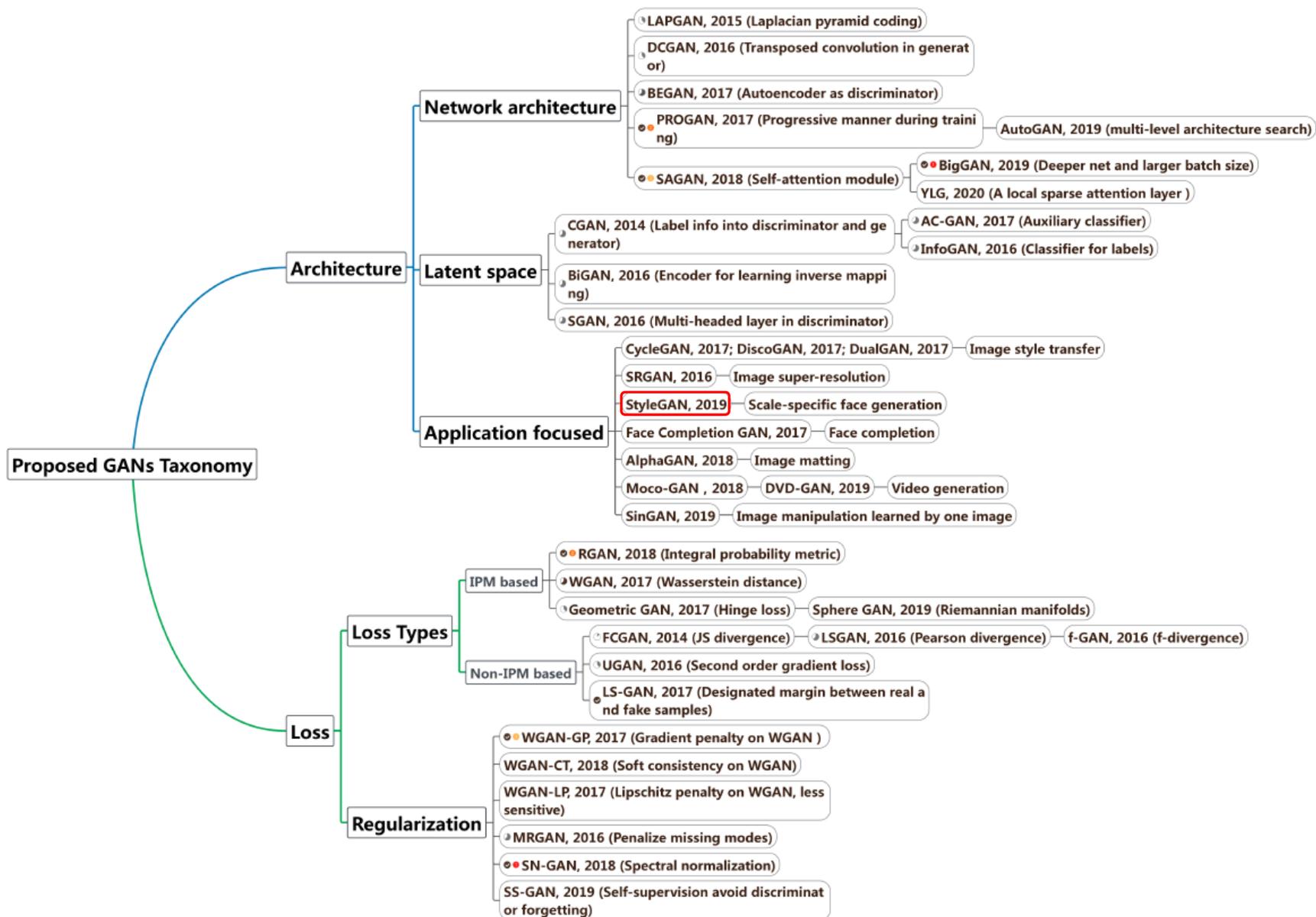
1. Avaliar os **tipos GANs** para utilização em sistema gerador de imagens preditivo, para determinar as possíveis idades dos desaparecidos;
2. Efetuar um **levantamento de banco de imagens** utilizados na literatura para o treino de modelos usando GANs;
3. Desenvolver ou **adaptar um modelo de GAN**, por meio de técnicas identificadas em publicações acadêmicas e pesquisas já realizadas.
4. **Avaliar a solução** proposta por meio da utilização de banco de imagens públicos (para validação do modelo com uso de métricas de aprendizagem de máquina) e também a aplicação de questionário online de pesquisa para o público.

1. Introdução
- 2. Fundamentação Teórica**
3. Trabalhos Correlatos
4. Solução Proposta
5. Avaliação Experimental
6. Considerações Finais e Trabalhos Futuros

- Classe da IA, pertencente a uma das especializações de modelos generativos de *Machine learning*.
- Arquitetura baseada em redes neurais.
- Gerador e Discriminador.

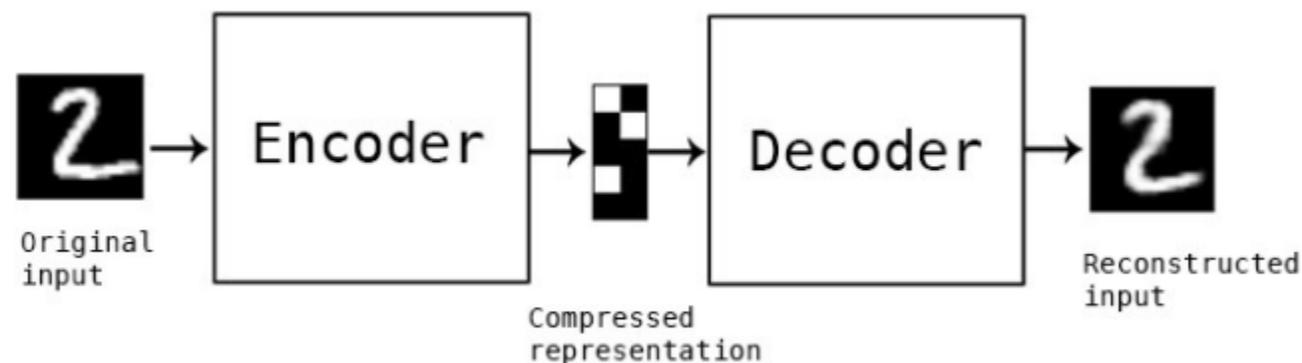


Fonte: (WU et al., 2017)



Fonte: (WANG et al., 2020)

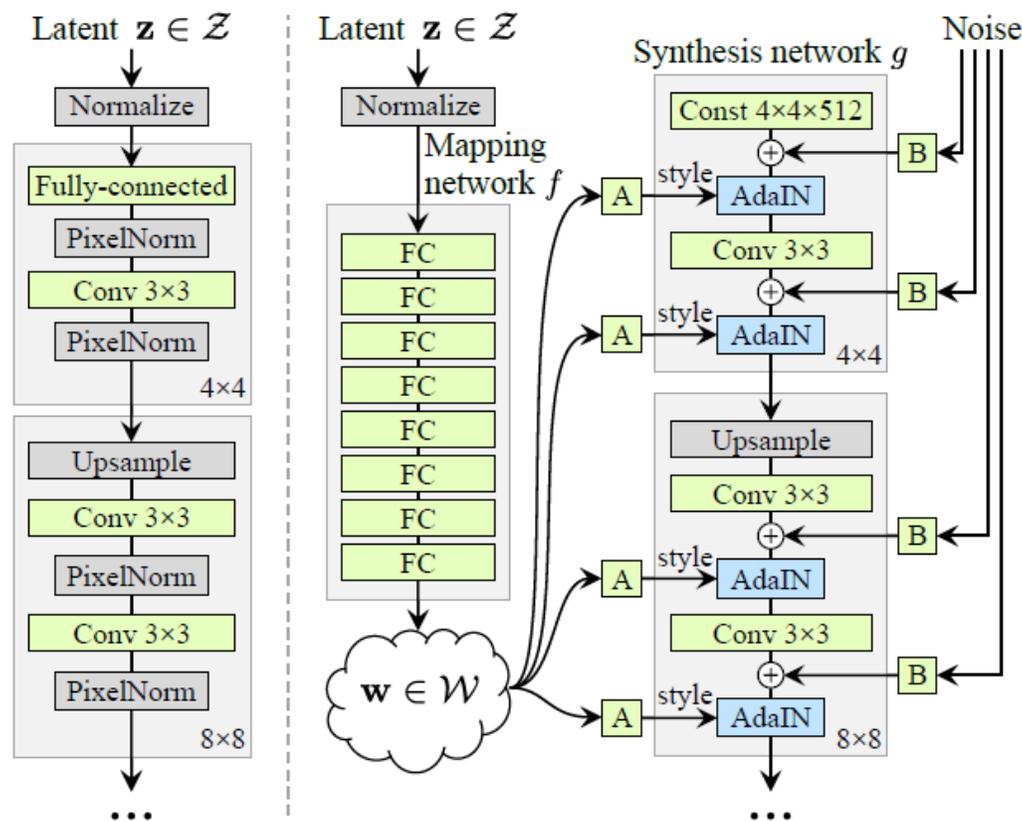
- Treinados para preservar o máximo de informações possíveis.
- Codificador (*Encoder*) e decodificador (*Decoder*).



Fonte: (BADR, 2019)<sup>1</sup>

<sup>1</sup> <https://towardsdatascience.com/auto-encoder-what-is-it-and-what-is-it-used-for-part-1-3e5c6f017726>

- A esquerda, gerador de um modelo **tradicional de GAN**.
- Diferenças: Baseada na **transferência de estilo**; Adição da **rede de mapeamento em 8 camadas** e vetor latente mapeado para o **espaço latente intermediário  $W$** .



Fonte: (KARRAS et al., 2019)

1. Introdução
2. Fundamentação Teórica
- 3. Trabalhos Correlatos**
4. Solução Proposta
5. Avaliação Experimental
6. Considerações Finais e Trabalhos Futuros

- **AgeGAN**, Face aging with conditional generative adversarial networks (ANTIPOV et. Al., 2017).
  - **Método baseado** em envelhecimento facial
  - **Objetivo**: Gerar ou retroceder um rosto
  - **Modelo baseado** em redes convolucionais
  - 53.2% Reconstrução inicial, 59.8% Com otimização *pixelwise* e 82.9% com otimização para preservação de identidade
  - **Limitação**: Modelo privado

- **Wavelets Based GAN**, Attribute-aware Face Aging with Wavelet-based Generative Adversarial Networks (LIU et. al., 2019).
  - **Método baseado** em envelhecimento facial
  - **Objetivo**: Manter os atributos faciais fidedignos e minimizar a ambiguidade da correspondência facial
  - **Modelo baseado** em *Wavelets*
  - 0.68 – Diferença entre idades de rostos sintéticos e não sintéticos; 98.26% na verificação facial; 96.07% gênero e 93.22% na preservação de raça. (MORPH)

- **InterfaceGAN**, Interpreting the Latent Space of GANs for Semantic Face Editing (SHEN et. al., 2020).
  - **Método baseado** em manipulações faciais de forma geral
  - **Objetivo**: Gerar e manipular faces através da edição semântica
  - **Modelo baseado** em transferência de estilo
  - 97.9% no conjunto de validação e 75.3% para todos os demais
  - **Limitação**: Tempo de **execução do modelo** (dependendo da quantidade de imagens a serem analisadas), é **demasiadamente longo**
  - **Arquitetura**: A mesma proposta (**StyleGAN**) por Karras et al. (2019)

- **StyleGAN**, A Style-Based Generator Architecture for Generative Adversarial Networks (KARRAS et. al., 2019).
  - **Método baseado** em manipulações de imagens de forma geral
  - **Objetivo**: Aumentar o desempenho obtido nesta tarefa
  - **Modelo baseado** em transferência de estilo
  - **FID, comprimento do caminho perceptivo e separabilidade linear** – Comprovaram que o **caminho completo e final são mais curtos** e se beneficiam melhor da **rede de mapeamento** no espaço  $W$
  - **Limitação**: Tempo de **treinamento da base de dados, aplicações de métricas e execução do modelo** (dependendo da quantidade de imagens a serem analisadas), são **demasiadamente longos**
    - Treinamento da base de dados: 5 semanas usando 1 GPU Tesla V100

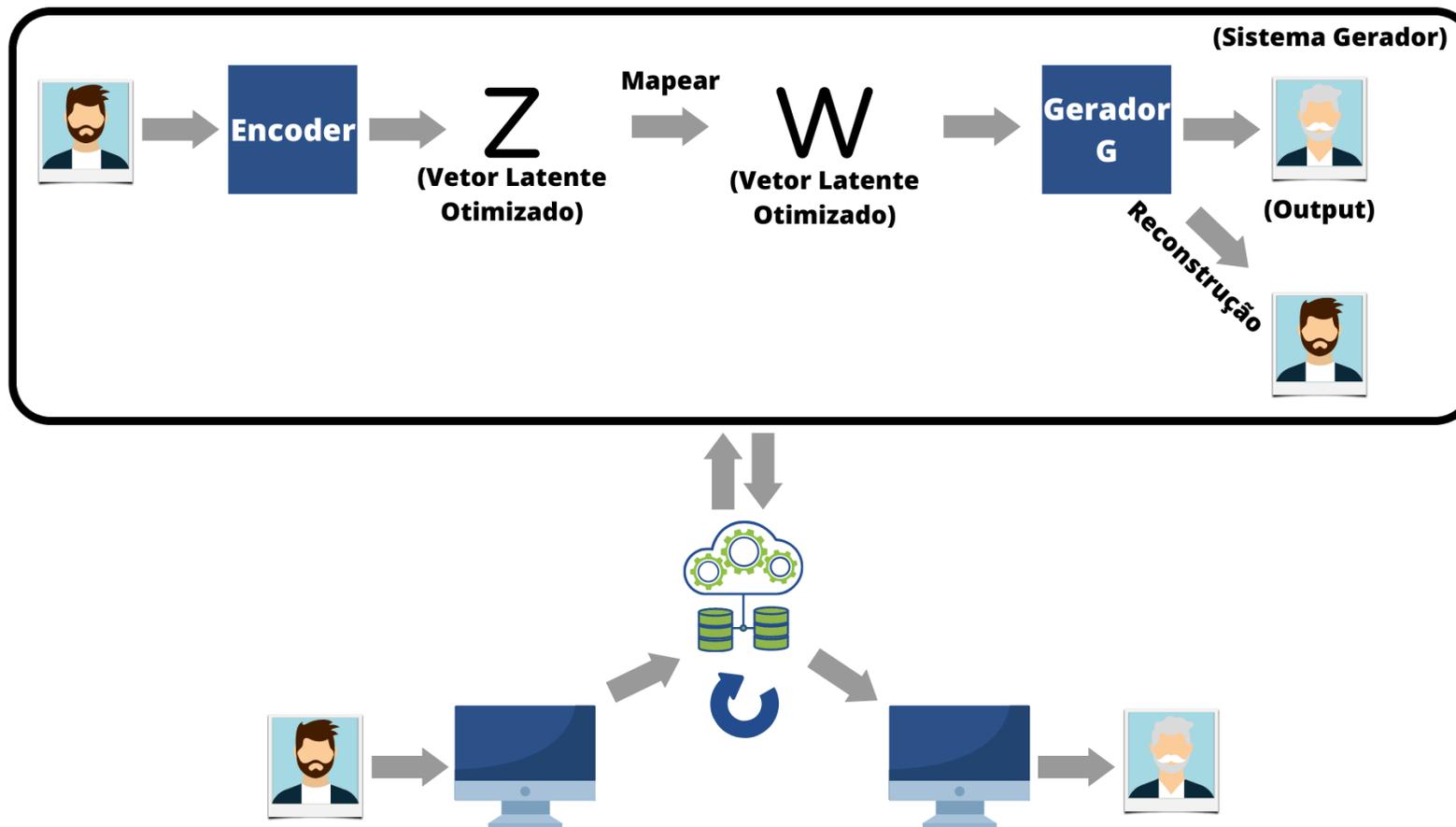
Tabela 7 – Classificação dos artigos por técnicas.

Artigos	Dados dos Trabalhos Correlatos					
	T.GAN	Dataset	Disp. Dataset	Tam. Dataset	Dim. Imgens	Result. de acurácia
Trabalho 1	Age-cGAN	IMDB-Wiki_cleaned	Público	120K	Não consta	82.9%
Trabalho 2	Wavelet-based GAN	MORPH e CACD	Privado	55K e 163K446	Não consta	98.26% e 98.44%
Trabalho 3	StyleGAN e PGGAN	CelebA e CelebA-HQ	Público	200k+ e 30K	178 x 218 e 1024 x 1024	97.9% e 75.3%
Trabalho 4	StyleGAN	CelebA-HQ e FFHQ	Público	30K e 70k	1024 x 1024	–

Fonte: Própria do autor.

1. Introdução
2. Fundamentação Teórica
3. Trabalhos Correlatos
- 4. Solução Proposta**
5. Avaliação Experimental
6. Considerações Finais e Trabalhos Futuros

Big picture do funcionamento do sistema proposto



Fonte: Própria do autor.

## Ferramentas e implementações

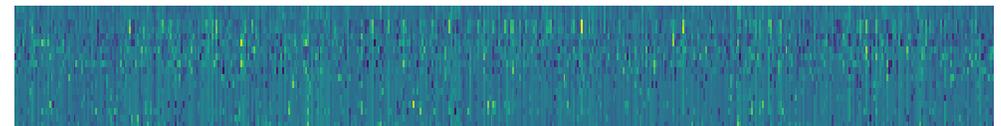
- Python (v3.7.12)
- Tensorflow (v2.4.1) – Compatibilidade com Tensorflow (v1.x)
- Keras (v2.3.1)
  - Prover suporte a implementação do modelo
  - Auxiliar nas etapas de construção de todas as redes que irão compor o modelo
  - *Encoder*, rede perceptual, gerador e discriminador

## Estrutura do GAN-AGE e o fluxo de execução

- Solução proposta dividida em duas partes:
  - **Stylegan-encoder e InterfaceGAN**
- **Stylegan-encoder**
  - Pré-processamento das imagens (redimensionamento, realinhamento e recorte)
  - Treino dos vetores latentes (através do encoder)

```
[[-0.11154527  0.16589175  0.05156069  ... -0.0316213  -0.00482604  
-0.1171001  ]  
[-0.06731246  0.05978122  0.00258037  ... -0.03100839  0.02400584  
-0.05867904]  
[ 0.04434378 -0.11088581 -0.11601208  ... -0.12592703  0.09350222  
-0.05128658]  
...  
[-0.00307101  0.05989887  0.00172584  ... -0.03182965 -0.22877331  
-0.05876277]  
[ 0.04461117  0.05988777  0.00165188  ... -0.03172294  0.02381369  
-0.05854049]  
[ 0.04404766  0.1320202  -0.04795876  ... -0.03156261  0.0248781  
-0.24220091]]
```

Fonte: Própria do autor.



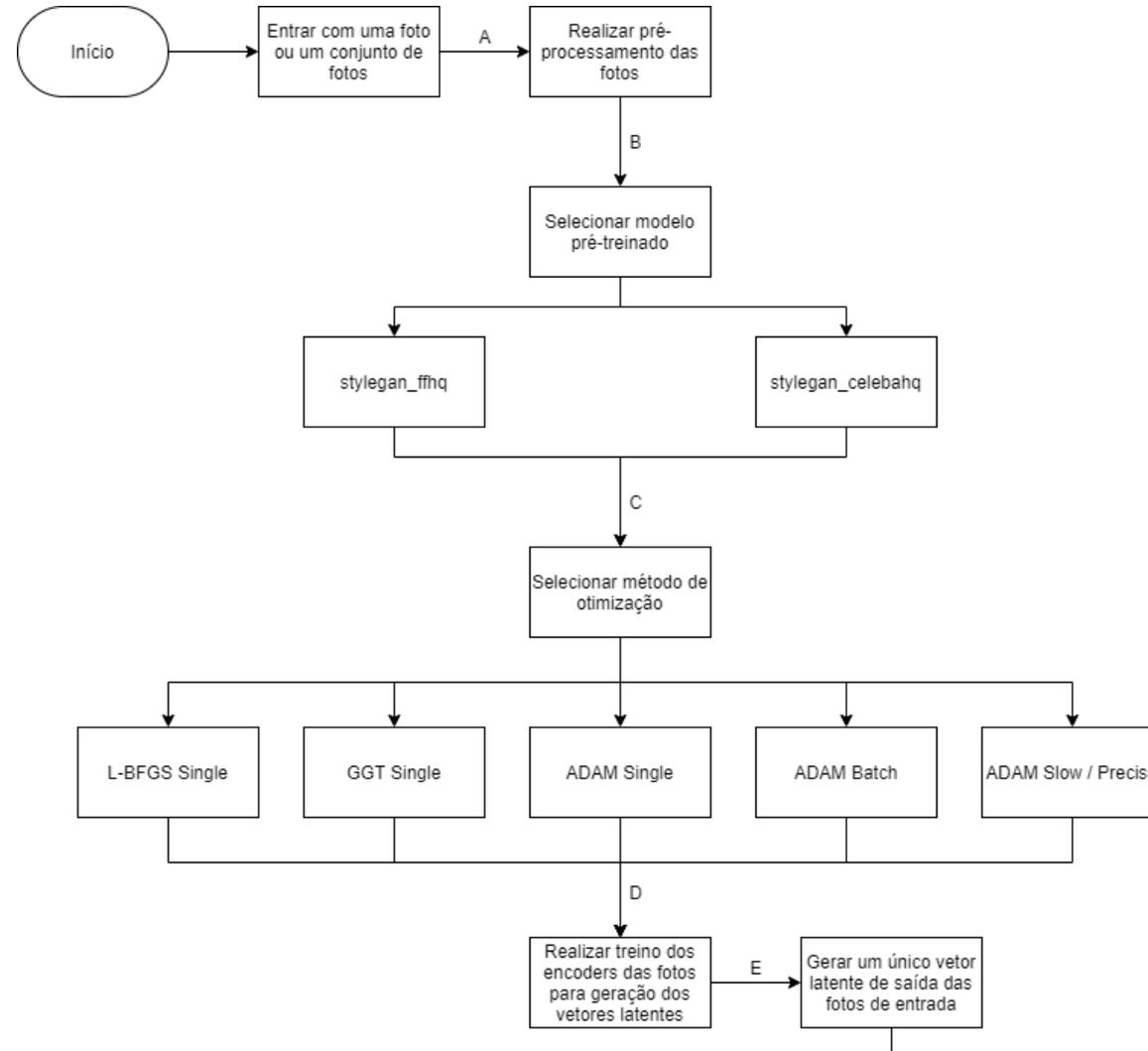
Fonte: Própria do autor.

## Estrutura do GAN-AGE e o fluxo de execução

### ○ InterfaceGAN

- Realiza a predição
- Parâmetros para o script de predição
  - **-m**, para especificar o modelo gerador
  - **-b**, para especificar o caminho dos limites semânticos
  - **-s**, para especificar o tipo de espaço latente
  - **-i**, para especificar o caminho dos códigos latentes de saída
  - **--start\_distance** e **--end\_distance** para especificar o intervalo de manipulação no espaço latente dos códigos latentes das imagens
  - **--steps**, para especificar o número de passos para cada edição sofrida pela imagem

## Fluxo de execução – Parte 1 (stylegan-encoder)



Fonte: Própria do autor.

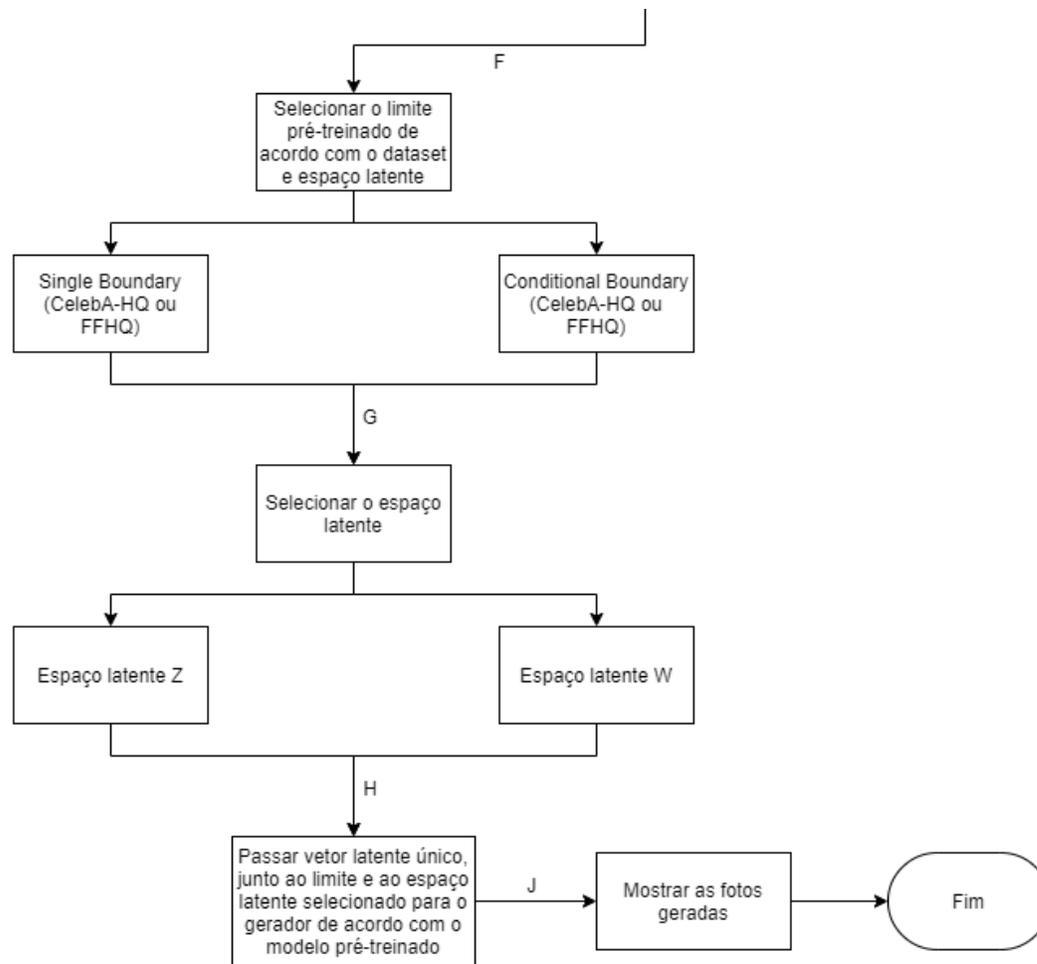
## Representação do vetor latente de saída único (output vector)

```
[[ -0.11154527  0.16589175  0.05156069 ... -0.0316213  -0.00482604  
  -0.1171001 ]  
 [ -0.06731246  0.05978122  0.00258037 ... -0.03100839  0.02400584  
  -0.05867904 ]  
 [  0.04434378 -0.11088581 -0.11601208 ... -0.12592703  0.09350222  
  -0.05128658 ]  
 ...  
 [ -0.00307101  0.05989887  0.00172584 ... -0.03182965 -0.22877331  
  -0.05876277 ]  
 [  0.04461117  0.05988777  0.00165188 ... -0.03172294  0.02381369  
  -0.05854049 ]  
 [  0.04404766  0.1320202  -0.04795876 ... -0.03156261  0.0248781  
  -0.24220091 ]]
```

```
[[ [ 3.5164185e-02  8.7448388e-01  4.1190456e-03 ... -3.2328263e-02  
    2.8902052e-02 -6.4246811e-02 ]  
 [ 4.2171717e-02  6.1537929e-02  1.6383012e-03 ... -1.4362061e-01  
    2.3639482e-02 -5.6101970e-02 ]  
 [ 2.7927639e-02  5.9848741e-02 -6.9916588e-03 ... -3.6133740e-02  
    2.0786921e-02 -7.0332840e-02 ]  
 ...  
 [ 8.2116053e-02  5.9428103e-02  5.1157642e-03 ... -5.8333933e-01  
   -3.6734837e-01 -5.9217859e-02 ]  
 [ 4.2782553e-02  5.9901588e-02  1.5321837e-03 ... -3.1028477e-02  
    2.7191881e-02 -6.1027352e-02 ]  
 [ 4.3684982e-02  6.0518138e-02  4.6519991e-03 ... -3.4881182e-02  
    2.5748447e-02 -5.8131471e-02 ]  
 ...  
 [ 3.8195748e-02  6.1177671e-02 -1.4239269e-03 ... -3.4417495e-02  
    2.5452184e-02 -5.3888798e-02 ]  
 [ 4.4860162e-02  1.7212820e-01 -5.4086518e-04 ... -3.8098142e-02  
    4.3797788e-01 -5.9588037e-02 ]  
 [ 2.9392673e-02  4.9178787e-02  1.3289428e-02 ... -4.0674064e-02  
    2.7627770e-02 -6.1357412e-02 ]  
 ...  
 [ 3.9399948e-02  5.9380323e-02 -2.0215327e-01 ... -3.1681620e-02  
    2.1798167e-02 -5.8346201e-02 ]  
 [ 4.5449503e-02  5.9373695e-02  2.3012604e-03 ... -3.3875611e-02  
    2.6413318e-02 -5.8669128e-02 ]  
 [ 4.4052366e-02  5.7439275e-02  6.0000038e-04 ... -2.7203199e-02  
    2.3466520e-02 -6.3467908e-01 ]]
```

Fonte: Própria do autor.

## Fluxo de execução – Parte 2 (InterfaceGAN)



Fluxo de execução da aplicação

Fonte: Própria do autor.

1. Introdução
2. Fundamentação Teórica
3. Trabalhos Correlatos
4. Solução Proposta
- 5. Avaliação Experimental**
6. Considerações Finais e Trabalhos Futuros

- Aplicando o método proposto (AGE-GAN)<sup>1</sup> sobre *benchmarks* públicos pré-treinados como FFHQ<sup>2</sup> e CelebA-HQ<sup>3</sup>.
- Experimentos foram conduzidos na plataforma de desenvolvimento Google Colab.
  - 12.72GB RAM
  - 68.40 Disco Temporário
  - GPU escolhida automaticamente
  - Só é Permitido escolher GPU na Versão Colab Pro
  - GPUs: Nvidia K80s, T4s, P4s e P100s

---

<sup>1</sup> <https://github.com/jpparreirap/TCC2>

<sup>2</sup> <https://github.com/NVLabs/stylegan>

<sup>3</sup> [https://github.com/tkarras/progressive\\_growing\\_of\\_gans](https://github.com/tkarras/progressive_growing_of_gans)

Investiga-se as seguintes questões de pesquisa:

1. O sistema proposto é capaz de gerar uma imagem de um rosto humano envelhecido dado uma idade alvo?
2. As métricas do modelo proposto são suficientes para provar a confiabilidade deste sistema?
3. O sistema proposto é capaz de se aproximar do perfil previsto de uma pessoa desaparecida com precisão, contribuindo nas buscas realizadas por investigadores?

Experimentos foram realizados com o autor deste trabalho e celebridades como Will Smith, Eunice Baía, Daniele Suzuki e Leonardo Dicaprio.

## Will Smith



Fonte: Própria do autor

## Eunice Baía



A) Input



B) Reconstrução Inicial



C) Output



D) Original



## Daniele Suzuki



A) Input



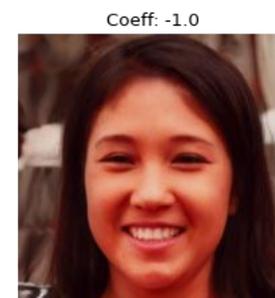
B) Reconstrução Inicial



C) Output



D) Original



Fonte: Própria do autor

## Leonardo Dicaprio



A) Input



B) Reconstrução Inicial



C) Output



D) Original



Coeff: -3.0



Coeff: -2.0



Coeff: -1.0



Coeff: 0.0

Fonte: Própria do autor

Métricas usadas para avaliação do modelo da GAN:

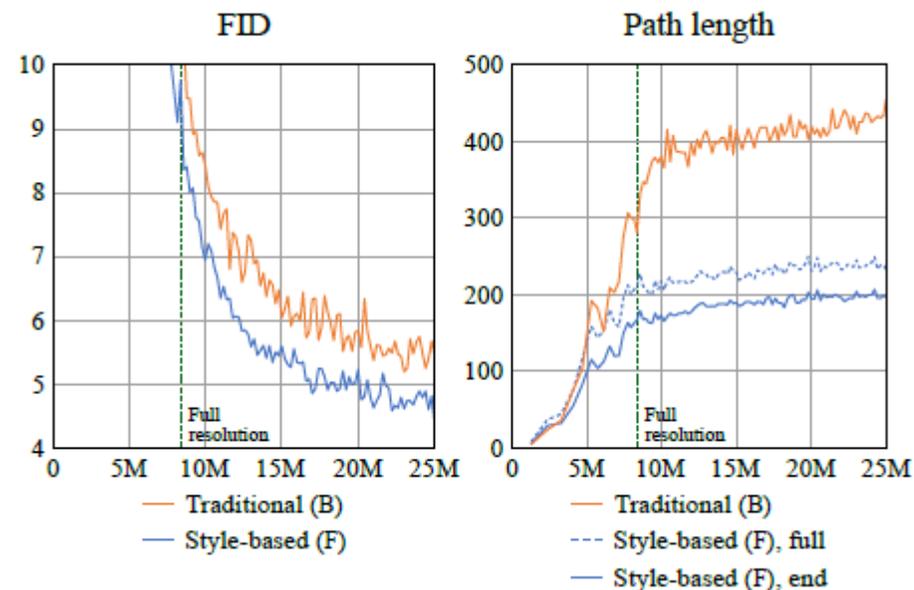
1. Fréchet Inception Distance (FID) – Menor melhor
  - **320.94** de qualidade da imagem, em relação às imagens geradas
2. Comprimento do caminho perceptivo e separabilidade linear

Espaço Latente	Comprimento do caminho		Separabilidade
	Completo	Final	
Z	670.7377	660.6735	44.3297
W	<b>235.3640</b>	<b>197.4768</b>	<b>1.3831</b>

Fonte: Própria do autor.

- Dados coletados a partir dos experimentos conduzidos por Karras et al. (2019).
- Base de dados usada: **FFHQ**, tamanho de **70 mil imagens**.
- Eixo horizontal, representa a quantidade de **imagens vista pelo discriminador**.
- Eixo vertical, representa o **resultado obtido**. Quanto menor melhor.

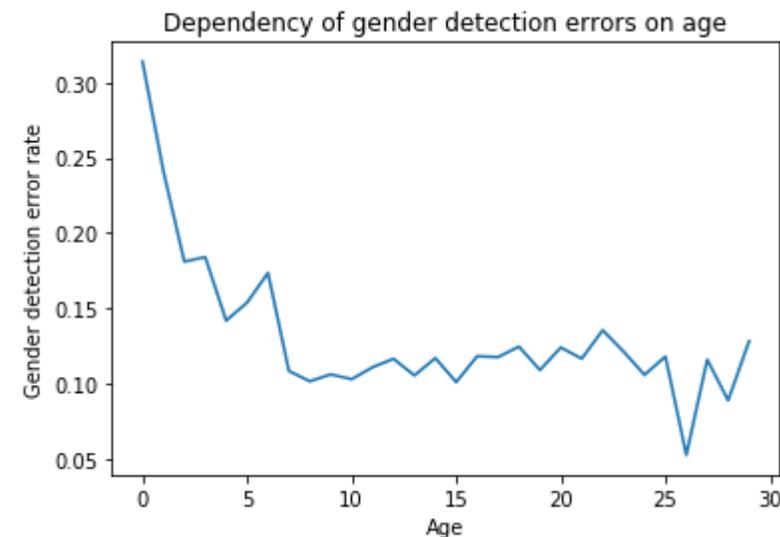
Curso que as métricas FID e PPL seguem (Dataset FFHQ)



Fonte: (KARRAS et al., 2019)

- Dados coletados a partir dos experimentos conduzidos por Nikitko (2019).
- Base de dados usada: **FFHQ**, tamanho de **70 mil imagens**.
- Eixo horizontal, representa a **idade das pessoas**.
- Eixo vertical, representa a **taxa de erro para a detecção de gênero**.

Relação entre a dependência da idade e os erros de detecção de gênero



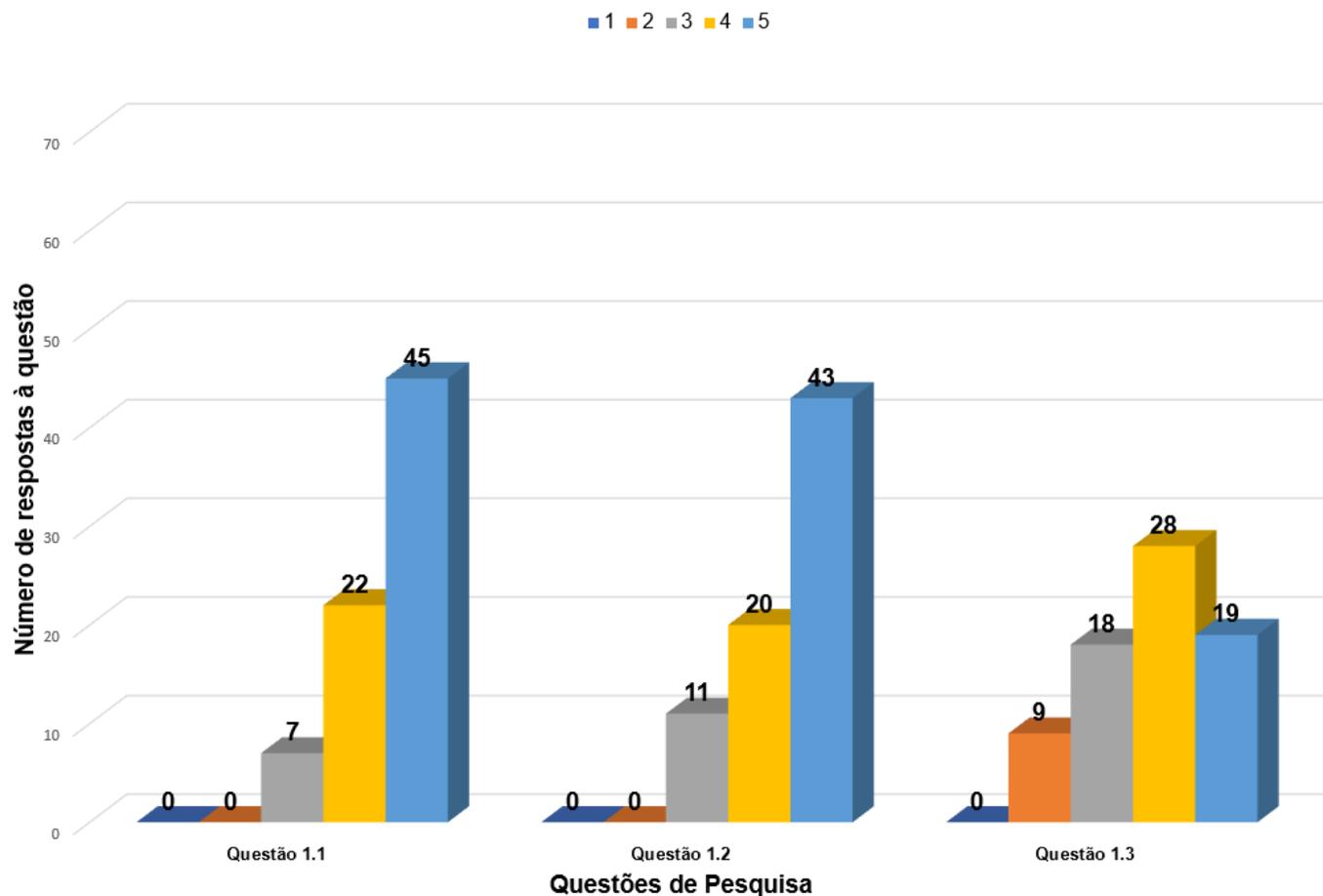
Fonte: (NIKITKO, 2019)

- Questionário online para avaliação dos resultados produzidos pela GAN:
- Alternativas de baseadas na **escala de Linkert**.
  - Alternativas de 1 à 5, onde 1 representa  **muito ruim** e 5  **muito bom**.
  - Plataforma utilizada para aplicar: **Google Forms**
  - Meios de publicação: **Whatsapp, Facebook e Instagram**
  - Foi permitido ao público o **compartilhamento do questionário**, visando **atingir o maior** número possível de **entrevistados**.
  - Número total de entrevistados que participaram da pesquisa foi de **74 pessoas**.

## Resultados de algumas questões selecionadas do questionário online

- Questão 1.1, nível de **nitidez**
- Questão 1.2, nível de **saturação**
- Questão 1.3, **semelhança** com os artistas
- Qualidade das imagens de forma geral

### Qualidade das imagens

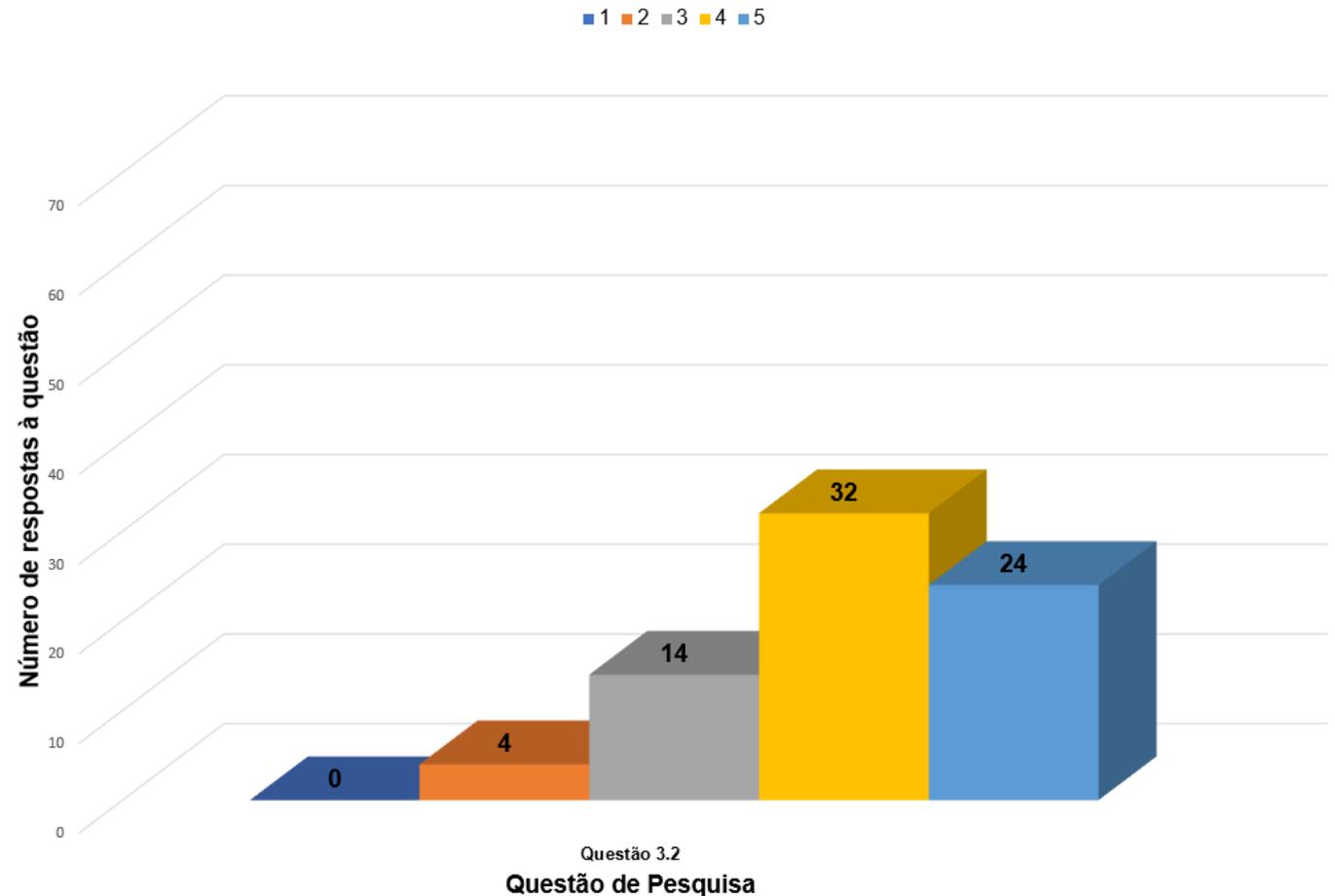


Fonte: Própria do autor

## Resultados de algumas questões selecionadas do questionário online

- Resultado pode ter sido determinado pela **previsão produzida** para a imagem de Will Smith.
- Pode estar relacionado à **falta de características bem definidas**.

### Qualidade da previsão de envelhecimento geral

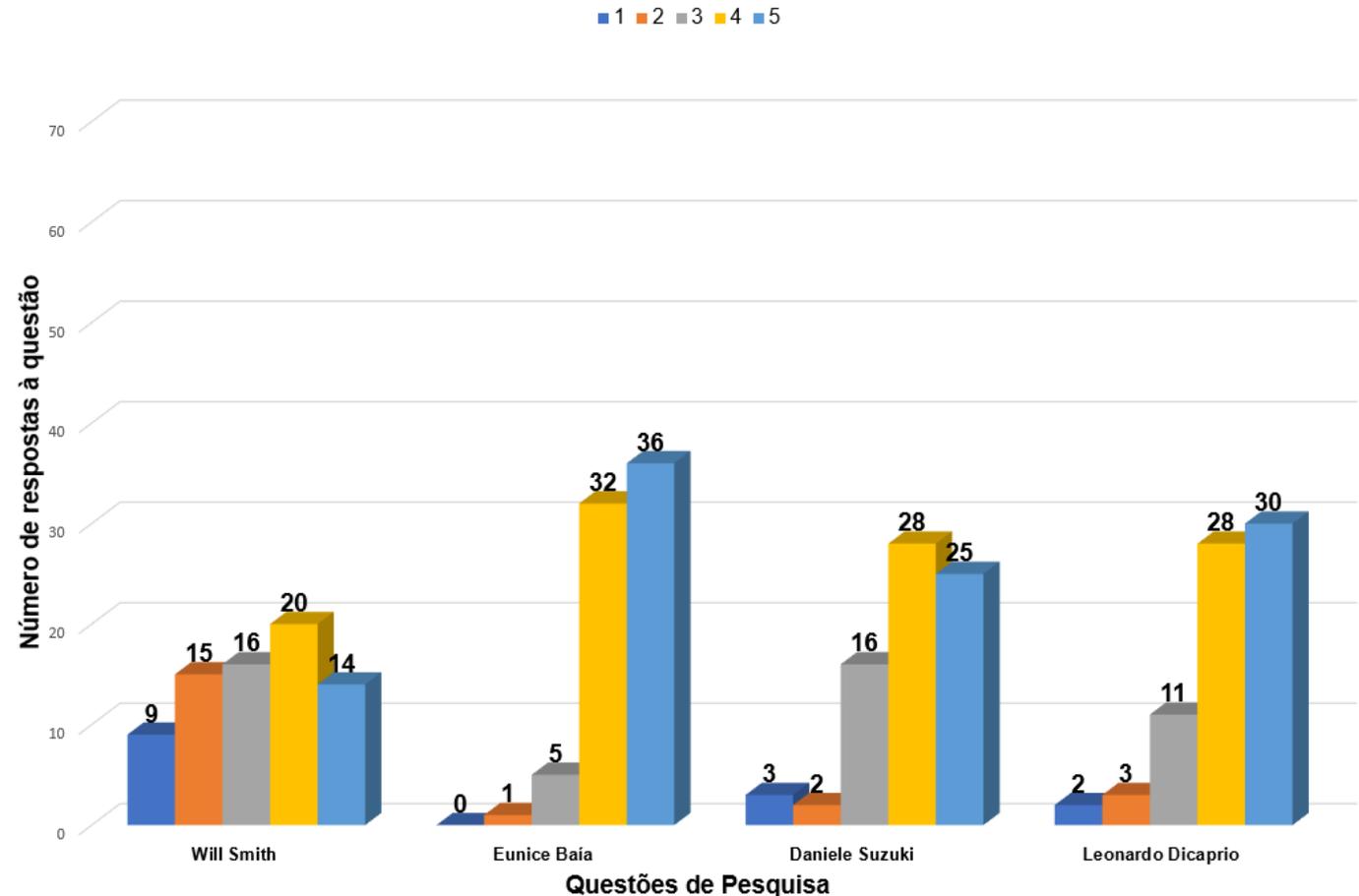


Fonte: Própria do autor

## Resultados de algumas questões selecionadas do questionário online

Qualidade da previsão de envelhecimento por celebridade

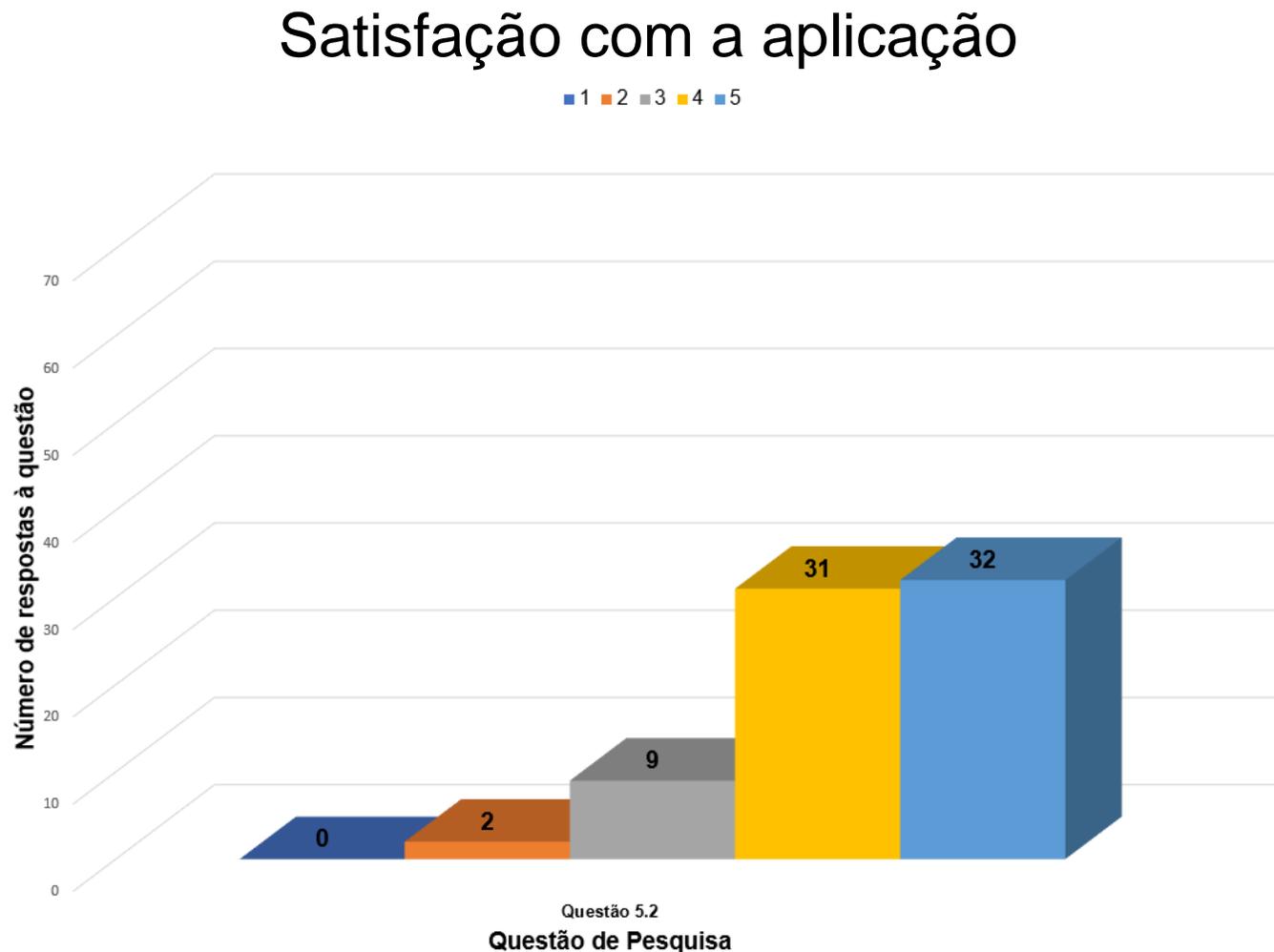
- Os resultados de Eunice Baía, foram **muito positivos**.
- Os resultados de Daniele Suzuki e Leonardo Dicaprio, foram **positivos porém não tanto quanto os de Eunice**.



Fonte: Própria do autor

## Resultados de algumas questões selecionadas do questionário online

- Os resultados com relação a satisfação da aplicação de **forma geral** foram **positivos**, ficando divididos entre **4** e **5**.



Fonte: Própria do autor

### Limitações da solução proposta:

- A **aplicação não é capaz** de prever **futuras características adicionais** que possam vir a surgir como: Cicatriz, procedimentos estéticos, cirúrgicos, uso de maquiagens, piercings, tatuagens, entre outros.
- **Pouca iluminação ou o excesso** pode afetar na previsão.
- Acessórios como óculos usados **no início da previsão**, podem afetar o resultado da previsão.
- **O brilho ou objetos espelhados** podem afetar drasticamente o resultado da previsão.

1. Introdução
2. Fundamentação Teórica
3. Trabalhos Correlatos
4. Solução Proposta
5. Avaliação Experimental
6. **Considerações Finais e Trabalhos Futuros**

- As métricas aplicadas mostraram que **o desempenho** ao se trabalhar com as imagens **no espaço latente  $W$  é superior** em relação **ao espaço  $Z$** , o que **corresponde** ao que foi descrito por **Karras et al. (2019)**.
- Os resultados obtidos do **na avaliação experimental** apontam **resultados promissores**, visto que a **aplicação é capaz de gerar e prever** um rosto humano a partir da **foto de entrada reconstruída**, com grande **aceitação pelo público** e pelas **métricas analisadas**.

- Sugestões de melhorias para trabalhos futuros:
  - **Incrementar a base de dados** com uma **maior variabilidade** de etnias, acessórios faciais, condições de iluminação e gênero
  - Adaptação da **técnica de preservação de identidade**, vista em (WANG et al., 2018), para StyleGAN
  - Adição de um classificador para **auxiliar na classificação de idade** das imagens geradas
  - Implementação de um **sistema Web interativo**, onde o investigador policial poderia realizar a entrada das fotos que se desejam prever

Obrigado pela atenção!

---

[jpparreirap@gmail.com](mailto:jpparreirap@gmail.com)