**FACULDADE DE TECNOLOGIA DE SÃO JOSÉ DOS CAMPOS**

**FATEC PROFESSOR Jessen Vidal**

**Leonardo Aleixo da Silva**

**Sistema de Identificação humana em câmeras de vigilância**

São José dos Campos

2022

**Leonardo Aleixo da Silva**

**Sistema de Identificação humana em câmeras de vigilância**

Trabalho de Graduação apresentado à Faculdade de Tecnologia de São José dos Campos, como parte dos requisitos necessários para a obtenção do título de Tecnólogo em Banco de Dados.

**Orientador: Carlos Augusto Lombardi Garcia, ME**

São José dos Campos

2022

Silva, Leonardo Aleixo

Sistema de Identificação humana em câmeras de vigilância.

São José dos Campos, 2022.

999f. (número total de folhas do TG)

Trabalho de Graduação – Curso de Tecnologia em Banco de dados

FATEC de São José dos Campos: Professor Jessen Vidal, 2022.

Orientador Interno ou Principal: Me, Carlos Augusto Lombardi Garcia.

1. Inteligência Artificial 1. 2. YOLO 2. 3. Câmera de segurança 3. I. Faculdade de Tecnologia. FATEC de São José dos Campos: Professor Jessen Vidal. Divisão de Informação e Documentação. II. Título

**REFERÊNCIA BIBLIOGRÁFICA**

Silva, Leonardo Aleixo. **Sistema de Identificação humana em câmeras de vigilância.** 2022. 999f. Trabalho de Graduação - FATEC de São José dos Campos: Professor Jessen Vidal.

**CESSÃO DE DIREITOS**

NOME(S) DO(S) AUTOR(ES): Leonardo Aleixo da Silva

TÍTULO DO TRABALHO: Sistema de Identificação humana em câmeras de vigilância.

TIPO DO TRABALHO/ANO: Trabalho de Graduação/2022.

É concedida à FATEC de São José dos Campos: Professor Jessen Vidal permissão para reproduzir cópias deste Trabalho e para emprestar ou vender cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte deste Trabalho pode ser reproduzida sem a autorização do autor.

|  |  |
| --- | --- |
| \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  Leonardo Aleixo da Silva  Avenida Pedro Friggi, 3100 – Bloco 12 Apto 34  12223-430, São José dos Campos – São Paulo |  |

**Leonardo Aleixo da Silva**

**Sistema de Identificação humana em câmeras de vigilância.**

Trabalho de Graduação apresentado à Faculdade de Tecnologia de São José dos Campos, como parte dos requisitos necessários para a obtenção do título de Tecnólogo em Banco de Dados.

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Titulação, Nome do Componente da Banca - Sigla da Instituição**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Titulação, Nome do Componente da Banca - Sigla da Instituição**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**ME, Carlos Lombardi Garcia**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Nome do Coorientador (se existir) - Sigla da Instituição**

**\_\_\_\_\_/\_\_\_\_\_/\_\_\_\_\_**

**DATA DA APROVAÇÃO (dia da banca)**

**AGRADECIMENTOS**

Agradeço ao professor e orientador Carlos Lombardi Garcia pelo apoio no desenvolvimento e instrução da utilização das tecnologias corretas, aos demais professores pelos conhecimentos transmitidos e aos colegas de faculdade e trabalho que incentivaram o desenvolvimento.

**RESUMO**

Apresentação concisa dos pontos relevantes do documento deve ser exposta no resumo. No presente caso o resumo será informativo, assim deverá ressaltar o objetivo, a metodologia, os resultados e as conclusões do documento. A ordem desses itens depende do tratamento que cada item recebe no documento original. O resumo deve ser composto por uma sequência de frases concisas, afirmativas e não em enumeração de tópicos. Deve ser escrita em parágrafo único e espaçamento de 1,5 linhas. A primeira frase deve ser significativa, explicando o tema principal do documento. Deve-se usar o verbo na voz ativa e na terceira pessoa do singular. Quanto a sua extensão, o resumo deve possuir de 150 a 500 palavras.

**Palavras-Chave**: Com um mínimo de 3 e no máximo 6 palavras, separadas entre si por ponto e vírgula “;” e finalizadas por ponto. As palavras-chave sãopalavras representativas do conteúdo do documento.

**ABSTRACT**

O abstract é o resumo da obra em língua estrangeira, que basicamente segue o mesmo conceito e as mesmas regras que o texto em português. Recomenda-se que para o texto do abstract o autor traduza a versão do resumo em português e faça, se necessário, os ajustes referentes à conversão dos idiomas. É importante observar que o título e texto NÃO DEVEM estar em itálico.

**Keywords**: Recomenda-se que o autor traduza para o inglês as Palavras-Chave em português e faça, se necessário, os ajustes referentes à conversão dos idiomas.

**LISTA DE FIGURAS**

[Figura 1 – Rede Neural Convolucional DeepLearningBook, 2021). 13](#_Toc99478877)

[Figura 2 - Fluxograma de processo de análise 15](#_Toc99478878)

[Figura 3 – Processo de análise em Python com OpenCV2 16](#_Toc99478879)

[Figura 4 - Resultado obtido após processamento de imagem 17](#_Toc99478880)

**LISTA DE ABREVIATURAS E SIGLAS**

CNN Convolutional Neural Network

CPU Central Processing Unit

IP Internet Protocol

**SUMÁRIO**

[1. INTRODUÇÃO 9](#_Toc99478710)

[1.1. Objetivos do Trabalho 9](#_Toc99478711)

[1.2. Conteúdo do Trabalho 9](#_Toc99478712)

[2. FUNDAMENTAÇÃO TÉCNICA 10](#_Toc99478713)

[2.1. Inteligência artificial 10](#_Toc99478714)

[2.2. Rede neural 10](#_Toc99478715)

[2.14 CNN – Convolutional Neural Network 13](#_Toc99478716)

[3. DESENVOLVIMENTO 14](#_Toc99478717)

[Neste capítulo será abordado o desenvolvimento do projeto. A primeira seção abordará a arquitetura, assim como os requisitos do projeto. A segunda seção será voltada às tecnologias e ferramentas utilizadas para tornar o projeto possível. 14](#_Toc99478718)

[3.1. Requisitos 14](#_Toc99478719)

[3.2. Requisitos Funcionais 14](#_Toc99478720)

[4. RESULTADOS 18](#_Toc99478721)

[4.1. Resultados Técnicos 18](#_Toc99478722)

[4.2. Aprendizagens 18](#_Toc99478723)

[5. CONSIDERAÇÕES FINAIS 19](#_Toc99478724)

[5.1. Contribuições 19](#_Toc99478725)

[5.2. Trabalho Futuros 19](#_Toc99478726)

[6. REFERÊNCIAS 20](#_Toc99478727)

# 1. INTRODUÇÃO

Muitas pessoas estão familiarizadas com o reconhecimento facial por usar a tecnologia de face ID para desbloqueio do celular. Essa tecnologia não é apenas um banco de dados contendo uma quantidade enorme de dados para validar o rosto e sim uma matemática capaz de calcular através dos traços do rosto suas características e distinguir se as informações são favoráveis ou não.

Essa tecnologia com um processamento de dados avançado possui outras funções como identificação de pessoas em câmera de segurança e com base no resultado foi definido um projeto para instalação das funções de identificação dentro de escolas públicas creches entre outros locais públicos.

## 1.1. Objetivos do Trabalho

O objetivo geral deste trabalho é desenvolver dois sistemas; O primeiro sendo capaz de configurar os ambientes de transferência de imagens e relatórios de capturas; O segundo seria capaz de identificar seres humanos em intervalos de tempo configurados no primeiro sistema.

.

Para a consecução deste objetivo foram estabelecidos os objetivos específicos:

* + Identificação das melhores tecnologias de detecção humana no mercado.
  + Criação de um gerenciamento de servidor FTP.
  + Criação de alerta por detecção humana.
  + Criação de um dashboard para detecção.

## 1.2. Conteúdo do Trabalho

O presente trabalho está estruturado em seis Capítulos, cujo conteúdo é sucintamente apresentado a seguir:

**Capítulo 1** Apresenta informações de introdução ao trabalho.

**Capítulo 2** Apresenta a fundamentação das tecnologias.

**Capítulo 3** Apresenta o desenvolvimento da solução.

**Capítulo 4** Apresenta os resultados obtidos e aprendizagens.

**Capítulo 5** Apresenta as considerações finais deste trabalho a partir da análise dos resultados obtidos.

**Capítulo 6** Apresenta as referências utilizadas para desenvolvimento.

# 2. FUNDAMENTAÇÃO TÉCNICA

Este capítulo tem por objetivo apresentar as tecnologias e termos utilizados para o entendimento da identificação humana por imagem.

O capítulo é subdivido em seções que descrevem o conceito de inteligência artificial.

Seções posteriores mostram uma breve explicação sobre as tecnologias, Sistema de controle de versão de dados (Github), Linguagens de programação Typescript, Java e Python. Os frameworks e bibliotecas YOLO, CNN (Convolutional Neural Network), Spring e Angular 2. Banco de dados utilizado Mysql.

## 2.1. Inteligência artificial

Esta tecnologia está relacionada ao processo e a capacidade de pensamento superpoderoso e a análise de dados para uma finalidade em específico, por exemplo, identificação de mudanças climáticas em imagens meteorológicas, projeções de lucros futuros, identificação de objetos entre outros.

(ORACLE, 2022)

## 2.2. Rede neural

Redes neurais são processamentos que refletem o pensamento humano, permitindo que programas de computador reconheçam padrões e resolvam problemas comuns nos campos de IA (Inteligência Artificial), usando machine learning e deep learning.

(IBM, 2022)

**2.3 Machine Learning**

É uma tecnologia onde os computadores têm a capacidade de aprender de acordo com as respostas esperadas por meio associações de diferentes dados, os quais podem ser imagens, números e tudo que essa tecnologia possa identificar. Machine Learning é o termo em inglês para a tecnologia conhecida no Brasil como aprendizado de máquina.

(IBM, 2022)

**2.4 Deep Learning**

É um tipo de machine learning que treina computadores para realizar tarefas como seres humanos, o que inclui reconhecimento de fala, identificação de imagem e previsões. Em vez de organizar os dados para serem executados através de equações predefinidas, o deep learning configura parâmetros básicos sobre os dados e treina o computador para aprender sozinho através do reconhecimento padrões em várias camadas de processamento.

(SAS, 2022)

**2.5 Git**

É um sistema de controle de versões distribuído, usado principalmente no desenvolvimento de software, mas pode ser usado para registrar o histórico de edições de qualquer tipo de arquivo

(Atlassian, 2022)

**2.6 Typescript**

 É uma linguagem de programação de código aberto desenvolvida pela Microsoft. É um superconjunto sintático estrito de JavaScript e adiciona tipagem estática opcional à linguagem.

(Microsoft, 2022)

**2.7 Python**

É uma linguagem de programação de alto nível, interpretada de script, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte

(Python Software Foundation, 1991)

**2.8 Java**

Java é uma linguagem de programação orientada a objetos desenvolvida na década de 90 por uma equipe de programadores chefiada por James Gosling, na empresa Sun Microsystems.

(Sun Microsystems, 1991)

**2.9 YOLO**

YOLO (You Only Look Once) é um método de detecção de objetos de passada única que utiliza uma rede neural convolucional como extrator de características.

(Joseph Redmon, 2015)

**2.10 Spring**

O Spring é um framework Java criado com o objetivo de facilitar o desenvolvimento de aplicações, explorando, para isso, os conceitos de Inversão de Controle e Injeção de Dependências.

(DevMedia, 2022)

**2.11 Angular 2**

É uma plataforma de desenvolvimento mobile e web desktop, baseada em TypeScript.

(Google, 2016)

**2.12 MySql**

O MySQL é um sistema de gerenciamento de banco de dados, que utiliza a linguagem SQL como interface.

(Oracle, 1995)

**2.13 OpenCV**

É uma biblioteca multiplataforma, totalmente livre ao uso acadêmico e comercial, para o desenvolvimento de aplicativos na área de Visão computacional, bastando seguir o modelo de licença BSD Intel.

(Intel, 2000)

# 2.14 CNN – Convolutional Neural Network

Uma Rede Neural Convolucional (ConvNet / Convolutional Neural Network / CNN) é um algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e vieses que podem ser aprendidos) a vários aspectos / objetos da imagem e ser capaz de diferenciar um do outro.

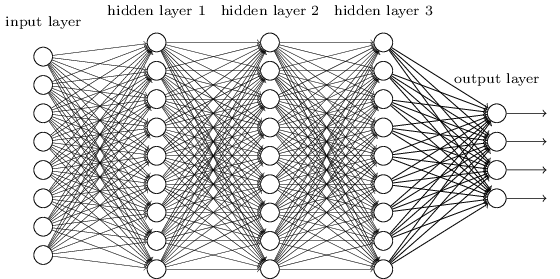


Figura 1 – Rede Neural Convolucional DeepLearningBook, 2021).

A figura 1 representa como funciona uma rede neural, onde se recebe uma entrada e vai passando por filtros distribuindo para outros filtros da imagem até chegar em um resultado. É como se estivesse resolvendo uma equação matemática seguindo as regras de ordem de cálculo.

# 3. DESENVOLVIMENTO

## Neste capítulo será abordado o desenvolvimento do projeto. A primeira seção abordará a arquitetura, assim como os requisitos do projeto. A segunda seção será voltada às tecnologias e ferramentas utilizadas para tornar o projeto possível.

## 3.1. Requisitos

Para que o sistema seja considerado pronto, é necessário atingir alguns requisitos.

## 3.2. Requisitos Funcionais

1. O sistema deve ser capaz de detectar um corpo humano através de uma leitura de imagem em diretório dinâmico.
2. O sistema deve enviar informações ao aplicativo de conversas telegram, a porcentagem de precisão detectada e as pessoas detectadas marcadas com um retângulo azul.

O sistema deve ser capaz de gerenciar as pastas onde será armazenado as imagens recebidas.

1. O sistema deverá autenticar um usuário para gerenciamento de módulos e permissões internas.
2. O sistema deverá ser capaz de criar usuários a partir de um usuário administrador geral.
3. O sistema deverá excluir as imagens que não detectar nenhum corpo humano.
4. O sistema deverá manter em uma pasta configurada as imagens processadas e depois do envio apagá-las.

**3.3 Requisitos Não Funcionais**

1. Portabilidade: O sistema deverá ser compilado e executado em qualquer plataforma.
2. O sistema deverá analisar quinze imagens por segundo.
3. O sistema deverá disponibilizar dados privados aos usuários.
4. O sistema deverá ter alta disponibilidade, cerca de 99% do tempo.

**3.4 Arquitetura do Sistema**

Diagrama

Descrição gerada automaticamente

Figura 2 - Fluxograma de processo de análise

A figura 1 mostra o fluxograma geral de como deverá funcionar o programa. O sistema, por sua vez, é encarregado de criar uma conexão FTP com as câmeras IP’s, e quando detectar um humano em uma imagem recebida e analisada, deverá ser enviado para o aplicativo telegram como alerta a foto e de qual câmera o dado se refere, caso não seja detectado, deverá descartar a imagem.

**3.5 FTP**

O sistema utiliza uma conexão FTP feita através do Filezila disponibilizando uma pasta de acesso.

**3.6 Análise da imagem**

A análise necessita de duas tecnologias para ser completada, a primeira será o método de detecção YOLO, ele é uma metodologia de IA convolucional que possui uma base de dados, a segunda será o OpenCV que irá utilizar esse método YOLO para fazer a detecção. O motivo do uso de OpenCV é devido a sua alta otimização em processadores INTEL e integração com CUDA (API de conexão com placas de vídeo NVIDIA).

A biblioteca OpenCV possui diversos algoritmos em C++ que viabilizam (OPENCV,

2013):

* Captura de imagens: é possível através da OpenCV acessar câmeras embutidas, USB ou até Câmeras IP, e através destas obter imagens (também chamado de quadros);
* Modificação e pré-processamento de imagens;
* Detecção de objetos: para o caso de faces humanas, é possível detectar a face frontal e perfis esquerdo e direito;
* Reconhecimento facial: a OpenCV possui uma classe dedicada a esta atividade, denominada FaceRecognizer, que faz a previsão de uma determinada face baseada em imagens armazenadas em banco de dados.

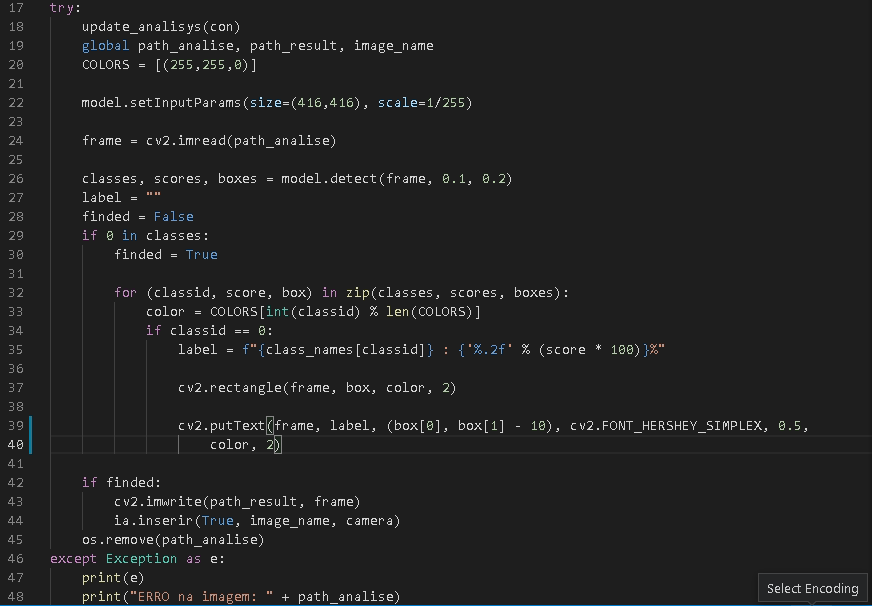


Figura 3 – Processo de análise em Python com OpenCV2

A figura 3 mostra como é feito o processo de análise da imagem em linguagem de programação.

O método recebe os parâmetros necessários para trabalhar com a imagem:

Linha 20 - Define qual será a cor utilizada no retângulo de marcação quando detectar uma pessoa.

Linha 22 - Representa a configuração dos parâmetros de escala da imagem, como por exemplo, altura e comprimento.

Linha 24 - Transforma a imagem em uma variável.

Linha 26 – Faz a detecção baseada no modelo definido (YOLO).

Linha 32 – Percorre os dados recuperados da detecção para fazer análise humana.

Linha 34 – Valida se existe ser humano encontrado na foto

Linha 35, 37 & 39 – Adicionam a imagem os dados de porcentagem, retângulo e texto descritivo.



Figura 4 - Resultado obtido após processamento de imagem

# 4. RESULTADOS

Para maior precisão nos resultados é necessário um número significativo de treinamentos para a rede neural, indicando tanto os valores positivos e negativos em relação a aprendizagem.

A Tecnologia YOLOv4 é uma rede bem treinada e disponibilizada gratuitamente para uso a qualquer programador.

Para a área de segurança a tecnologia de identificação humana com redes neurais são altamente recomendáveis.

## Resultados Técnicos

O processamento final chegou a ser 15 imagens por segundo em um ambiente com 8 gigabyte de memória ram e um CPU I5 7600.

O processamento final chegou a ser 25 imagens por segundo em um ambiente com 8 gigabyte de memória ram e uma placa de vídeo GFORCE 750 TI.

## 4.2. Aprendizagens

Toda rede neural convolucional deve ser treinada com muitos dados de forma assertiva aos pontos falsos e verdadeiros.

OpenCV é uma ótima biblioteca para utilização das redes neurais YOLO além de fornecer muitas funcionalidades de tratamento de imagens.

# 5. CONSIDERAÇÕES FINAIS

Este capítulo visa apresentar algumas das considerações a respeito do

desenvolvimento do trabalho, experiências, conclusões, assim como sugestões para trabalhos futuros.

## 5.1. Contribuições

Nessa seção deverão ser listadas as contribuições do trabalho, experiências e dificuldades dos autor no decorrer do trabalho.

## 5.2. Trabalho Futuros

Para trabalhos futuros, sugere-se a otimização do processamento de imagens com implementação de novas redes neurais ou novos treinamentos para a rede neural, como por exemplo YOLOv5.

# 6. REFERÊNCIAS

ORACLE. **Apresentação e Definição sobre o que é Inteligência Artificial.**   
Disponível em   
https://www.oracle.com/br/artificial-intelligence/what-is-ai/#:~:text=Em%20termos%20mais%20simples%2C%20a,base%20nas%20informa%C3%A7%C3%B5es%20que%20coletam.

IBM. **Apresentação e Definição sobre o que são as Redes Neurais.**Disponível emhttps://www.ibm.com/br-pt/cloud/learn/neural-networks#:~:text=As%20redes%20neurais%20refletem%20o,machine%20learning%20e%20deep%20learning.

IBM. **Apresentação e Definição sobre o que é Machine Learning.**  
Disponível em  
https://www.ibm.com/br-pt/analytics/machine-learningSAS. **Apresentação e Definição sobre o que é Deep Learning.**  
Disponível em  
https://www.sas.com/pt\_br/insights/analytics/deep-learning.html

YOLO. **Apresentação e Definição da metodologia YOLO.**  
Disponível em   
https://iaexpert.academy/2020/10/13/deteccao-de-objetos-com-yolo-uma-abordagem-moderna/