

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA

FACULDADE DE TECNOLOGIA DE INDAIATUBA

DR. ARCHIMEDES LAMOGLIA

CURSO DE TECNOLOGIA EM

ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

BIANCA CAVALCANTE BASSO

**INTELIGÊNCIA ARTIFICIAL PARA RECONHECIMENTO
DE EMOÇÕES EM UM AMBIENTE EMPRESARIAL**

INDAIATUBA
2019

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA
FACULDADE DE TECNOLOGIA DE INDAIATUBA
DR. ARCHIMEDES LAMOGLIA

CURSO DE TECNOLOGIA EM
ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

BIANCA CAVALCANTE BASSO

**INTELIGÊNCIA ARTIFICIAL PARA RECONHECIMENTO
DE EMOÇÕES EM UM AMBIENTE EMPRESARIAL**

Trabalho de Graduação apresentado por Bianca Cavalcante Basso como pré requisito para a conclusão do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas, da Faculdade de Tecnologia de Indaiatuba, elaborado sob a orientação da Prof.^a Dr.^a Maria das Graças Junqueira Machado Tomazela.

INDAIATUBA
2019

CENTRO ESTADUAL DE EDUCAÇÃO TECNOLÓGICA PAULA SOUZA

FACULDADE DE TECNOLOGIA DE INDAIATUBA

CURSO DE TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE

SISTEMAS

BIANCA CAVALCANTE BASSO

Banca Avaliadora:

Profa. Dra. Maria das Graças J. M. Tomazela	Orientadora
Prof. Dr. Aldo do Nascimento Pontes	Avaliador interno
Prof. Dr. Dilemar Piva Júnior	Avaliador externo – Fatec Itu

Data da defesa 09/12/2019

Dedico este trabalho, assim como as conquistas nessa
faculdade, à Deus, aos meus pais, à minha orientadora e aos
meus amigos.

AGRADECIMENTOS

Agradeço primeiramente a Deus, pois é graças a Ele que estamos todos aqui seguindo nossas vidas.

Agradeço meus pais, Andréia e Alvaro, pois sem eles eu não estaria aqui e não teria enfrentado uma segunda graduação superior.

Agradeço em especial à minha orientadora Dra. Graça, que durante todo o processo de desenvolvimento deste trabalho me auxiliou, me deu os devidos puxões de orelha e me ajudou a concluí-lo. Agradeço por não ter desistido da minha pessoa e por ter sempre mantido a paciência com o meu trabalho sangrento.

Agradeço aos meus amigos que me incentivaram a continuar e a perseverar nessa etapa de minha vida.

E por fim, agradeço aos meus colegas de trabalho que me ajudaram a testar a viabilidade da aplicação e em especial agradecer ao meu colega de trabalho João Paulo que me ajudou a realizar algumas alterações na aplicação da rede neural.

“Seja um parâmetro de qualidade. Algumas pessoas não
estão acostumadas a um ambiente onde a excelência é
esperada.”
Steve Jobs

RESUMO

Na rotina de trabalho das pessoas, repara-se que o ser humano precisa tomar uma série de decisões, além de buscar soluções para problemas propostos e impostos a ele, tudo isso com o máximo de rapidez. Pode influenciar direta ou indiretamente na vida do ser humano, ocasionando assim a pressão decorrente do dia a dia laboral. Essas mudanças podem causar um mau desempenho decorrente da insatisfação do colaborador, que às vezes não é percebida pelos gestores. Nesse contexto, surgiu a ideia de utilizar algoritmos de inteligência artificial e rede neural convolucional para armazenar e reconhecer expressões faciais para interpretação de emoções dos colaboradores por meio de seus rostos (em tempo real construindo dessa maneira, um histórico de emoções obtidas dos seus colaboradores). Para tanto, a realização deste estudo abordou os seguintes conceitos chave, que sustentam a pesquisa: Inteligência artificial, redes neurais artificiais, reconhecimento facial e inteligência emocional. Também foram exploradas outras pesquisas acadêmicas relacionadas, para auxiliar na idealização do tema proposto, além de verificar ferramentas, softwares e meios de implementação que poderiam auxiliar no desenvolvimento deste trabalho. Para a realização deste trabalho foi feita uma pesquisa experimental, que resultou em uma aplicação de reconhecimento facial (raiva, nojo, medo, feliz, triste, surpresa, neutro), utilizando Rede Neural Convolucional, além de ferramentas e bibliotecas de apoio para a utilização do modelo proposto. A acurácia do modelo foi melhor para reconhecimento de expressões neutras, com valores variando entre 95% e 99%; felizes, com acurácia variando entre 58% e 61%; e raiva, com acurácia variando entre 80% e 90%, enquanto as demais emoções resultaram em muitos falsos positivos.

Palavras chave: Inteligência Artificial, Reconhecimento Facial, Redes Neurais, Expressões Faciais, Gestão de Pessoas.

FIGURAS

Figura 1: Modelo do neurônio artificial proposto por McCulloch e Walter Pitts	18
Figura 2: Arquitetura de uma rede neural	19
Figura 3: Típico diagrama em bloco de uma Rede Neural Convolucional.....	23
Figura 4: Pontos na face	25
Figura 5: Arquitetura do projeto.....	34
Figura 6: Treinamento da rede neural	37
Figura 7: Reconhecimento da face e expressão: Raiva	38
Figura 8: Reconhecimento da face e expressão: Feliz	38
Figura 9: Reconhecimento da face e expressão: Neutra	38
Figura 10: Reconhecimento da face e expressão: Triste	39
Figura 11: Reconhecimento da face e expressão: Assustado	39
Figura 12: Log do sistema.....	39
Figura 13: Log do sistema com o <i>Frame</i> no ambiente de trabalho, colaborador 1.....	40
Figura 14: Log do sistema com <i>Frame</i> no ambiente de trabalho, colaborador 2.....	41
Figura 15: Log do sistema com <i>Frame</i> no ambiente de trabalho, colaborador 3.....	41
Figura 16: Colaborador 1 - Emoção Raiva.....	42
Figura 17: Colaborador 1 - Emoção Feliz.....	42
Figura 18: Colaborador 1 - Emoção Neutra	43
Figura 19: Colaborador 1 - Emoção Triste.....	43
Figura 20: Colaborador 1 - Emoção Assustado	44
Figura 21: Colaborador 2 - Emoção Raiva.....	44
Figura 22: Colaborador 2 - Emoção Feliz.....	44
Figura 23: Colaborador 2 - Emoção Neutro.....	45
Figura 24: Colaborador 2 - Emoção Triste.....	45
Figura 25: Colaborador 2 - Emoção Assustado	46
Figura 26: Colaborador 3 - Emoção Raiva.....	46
Figura 27: Colaborador 3 - Emoção Feliz.....	47
Figura 28: Colaborador 3 - Emoção Neutro.....	47
Figura 29: Colaborador 3 - Emoção Triste.....	47
Figura 30: Colaborador 3 - Emoção Assustado	48

TABELAS

Tabela 1 - Matriz de Confusão de um Classificador - para K classes	21
Tabela 2 - Matriz De Confusão - para 2 classes	21

QUADROS

Quadro 1: Divisão da Inteligência Artificial	12
Quadro 2: Trabalhos relacionados	27
Quadro 3: Sintetização dos resultados	44

SUMÁRIO

INTRODUÇÃO

CAPÍTULO I 15

Fundamentação Teórica 15

1.1 Conceitos chave 15

1.1.1 Inteligência Artificial 15

1.1.2 Redes Neurais Artificiais 18

1.1.2.1 Medidas de Desempenho dos Classificadores 20

1.1.2.2 Deep Learning 22

1.1.2.3 Redes Neurais Convolucionais 23

1.1.3 Reconhecimento facial 24

1.1.4 Inteligência emocional 26

1.2 Trabalhos relacionados 27

CAPÍTULO II 32

Metodologia 32

2.1 Natureza da Pesquisa 32

2.2 Variáveis de análise 32

2.3 Ferramentas Utilizadas 32

2.4 Experimento de Pesquisa 33

2.5 Critérios para avaliação da pesquisa 34

CAPÍTULO III 35

Apresentação, implementação, testes e resultados 35

3.1 Base de dados utilizada 35

3.2 Processo de treinamento da rede neural convolucional 36

3.3 Realização dos testes da aplicação 37

3.4 Análise e apresentação dos resultados 40

CONSIDERAÇÕES FINAIS 49

REFERÊNCIAS	51
ANEXOS	515

INTRODUÇÃO

Na rotina comum do ser humano, é possível verificar alterações nas expressões e emoções de acordo com os acontecimentos decorrentes do dia a dia. Algumas emoções são alteradas dependendo da situação a que o ser humano é exposto, podendo variar perante uma simples mensagem recebida, uma conversa, situações vistas, dentre outras diversas situações existentes. As emoções podem variar de diferentes formas, sendo as mais expressivas aos olhos do ser humano: a felicidade, tristeza, surpresa, raiva (BUSSO et al., 2004; KELTNER et al., 2003 apud ZANELATO et al, 2017).

Segundo Bergamini (2010 p.17), as pessoas se comportam de maneira diferente dentro das organizações, ou seja, “é evidente [...] o aspecto das diferenças de comportamento”. A mesma autora descreve como esses determinantes afetam o desenvolvimento do indivíduo em uma empresa, afirmando a existência de duas variáveis, variáveis individuais, que diz respeito a sua infância e adolescência e a fase adulta; e as variáveis ambientais, caracterizam-se pela forma que o ser humano se adapta ao ambiente social, cultural, os fatores físicos e muitos outros.

O ambiente empresarial envolve fatores internos como também externos, nos quais observa-se a disputa acirrada por clientes e lucros, o que faz com que os gestores requeiram excessivo empenho e produtividade de seus colaboradores, nesse processo não considerar a pressão que geram sobre os colaboradores. É essa condição que viabiliza este projeto, que a possibilidade de verificar as emoções dos colaboradores no meio empresarial, auxiliando supervisores, gestores a terem uma atenção a mais com seus colaboradores, sendo notificados pela aplicação.

Sistemas que usam Inteligência Artificial podem auxiliar os humanos a serem mais precisos em suas decisões, realizar funções com menos chances de errar, desenvolver sistemas que aprendem com o passado e tomar decisões com base nesses fatos passados, para ter uma assertividade maior.

O desenvolvimento de sistemas desse tipo serve para a realização de processos complexos, fazer com que máquinas realizem até mesmo tarefas perigosas e cansativas evitando o desgaste humano e a periculosidade de certas atividades.

A partir desse contexto, esta pesquisa busca responder à seguinte pergunta:

- Como algoritmos de Inteligência Artificial (IA) podem auxiliar na identificação das

expressões faciais em um ambiente empresarial e gerar um relatório sobre o estado aparente de seus colaboradores?

Assim, o objetivo deste trabalho foi utilizar algoritmos de inteligência artificial e rede neural convolucional para armazenar e reconhecer expressões faciais para interpretar emoções de colaboradores por meio de suas expressões faciais, em tempo real, facilitando assim a tomada de decisão nas avaliações e possíveis intervenções e encaminhamento desses colaboradores para profissionais especializados.

Como hipótese, caso seja aplicado uma ferramenta que possa auxiliar na identificação de expressões faciais, será possível alertar antecipadamente superiores para tomadas de decisões com relação ao indivíduo identificado.

Como método de pesquisa, este trabalho, utilizou o método experimental que segundo Gil (2002), consiste em determinar um objeto de estudo, selecionar as variáveis que seriam capazes de influenciá-lo e definir as formas de controle e de observação dos efeitos que a variável produz no objeto.

Com relação à estrutura da pesquisa, está estruturada como se segue: no Capítulo I, é exposta a Fundamentação Teórica, que traz os conceitos chave e os trabalhos relacionados que auxiliaram no desenvolvimento desta pesquisa. No capítulo II, é apresentada a Metodologia adotada neste trabalho, o percurso para a realização desta pesquisa experimental. No capítulo III, é apresentada a utilização da aplicação, a base de dados e os resultados obtidos por meio desta pesquisa. Por fim, são apresentadas as considerações finais e as referências.

CAPÍTULO I

Fundamentação Teórica

1.1 Conceitos chave

No embasamento desta pesquisa, optou-se por organizar este capítulo em duas partes. Primeiramente apresentam-se os conceitos chave que referenciam o trabalho, sendo eles: *inteligência artificial*, *redes neurais*, *reconhecimento facial* e *inteligência emocional*. Na segunda parte, apresenta-se um conjunto de trabalhos relacionados a esta pesquisa, decorrentes de estudos realizados na última década.

1.1.1 Inteligência Artificial

A Inteligência Artificial (IA) tem por finalidade criar sistemas inteligentes que possam executar tarefas de forma semelhante ao comportamento humano, é uma área que foi influenciada por diversas áreas como filosofia, matemática, economia, neurociência, psicologia, engenharia de computadores e linguística. As primeiras ideias sobre IA surgiram na década de 50, naquela época Turing propôs um teste para avaliar o nível de inteligência do computador. Esse teste se baseava na comparação das respostas de um computador e de um ser humano às indagações de uma pessoa, caso não houvesse uma forma de distingui-los, significaria que o computador era uma máquina inteligente (RUSSEL e NORVIG, 2003).

Segundo Russel e Norvig (2003) a IA surgiu para auxiliar na busca pelo conhecimento e a forma de se pensar, compreender, prever e manipular um mundo muito maior e mais complicado do que se imagina.

De acordo com Russel e Norvig (2003), para conceituar Inteligência Artificial, fazem essa divisão em quatro categorias: 1) pensando como um humano; 2) pensando racionalmente; 3) agindo como seres humanos e; 4) agindo racionalmente, conforme Quadro 1, é possível visualizar oito definições de IA, dispostas ao longo de duas dimensões. Em linhas gerais, as que estão na parte superior d quadro se relacionam a processos de pensamento e raciocínio, enquanto as definições da parte inferior se referem ao comportamento.

Para explicar melhor a Quadro 1, as definições do lado esquerdo medem o sucesso em termos de fidelidade ao desempenho humano, ou seja, quando se refere à definição da inteligência artificial baseando-se no ser humano, suas ações, a forma de agir, e sua forma de pensar como um todo. Enquanto as definições do lado direito medem o sucesso que em sua definição, se dá quando refere-se a parte racional, desfocando da parte emocional ou sentimental do ser humano. O sistema mede como sucesso quando a máquina apresenta uma resposta inteligente em uma determinada situação. Dessa maneira, mede o sucesso comparando-o com um conceito ideal de inteligência, chamado de racionalidade.

Quadro 1: Divisão da Inteligência Artificial

Pensando como um humano	Pensando racionalmente
<p>“O novo e interessante esforço para fazer os computadores pensarem (...) <i>máquinas com mentes</i>, no sentido total e literal.” (Haugeland, 1985)</p> <p>“[Automatização de] atividades que associamos ao pensamento humano, atividades como a tomada de decisões, a resolução de problemas, o aprendizado...” (Bellman, 1978)</p>	<p>“O estudo das faculdades mentais pelo uso de modelos computacionais.” (Charniak e McDermott, 1985)</p> <p>“O estudo das computações que tornam possível perceber, raciocinar e agir.” (Winston, 1992)</p>
Agindo como seres humanos	Agindo racionalmente
<p>“A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (Kurzweil, 1990)</p> <p>“O estudo de como os computadores podem fazer tarefas que hoje são melhor desempenhadas pelas pessoas.” (Rich and Knight, 1991)</p>	<p>“Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (Poole <i>et al.</i>, 1998)</p> <p>“AL... está relacionada a um desempenho inteligente de artefatos.” (Nilsson, 1998)</p>

Fonte: Russell e Norvig 2003. Pág. 3

Após compreender as divisões que deram início ao conceito de Inteligência Artificial, é necessário conhecer alguns dos principais paradigmas de IA, são eles: Sistemas Especialistas (SE), Raciocínio Baseado em Casos (RBC), Algoritmos Genéticos (AG), Lógica Fuzzy e Redes Neurais Artificiais (RNA). Tais paradigmas serão apresentados resumidamente a seguir.

Gonçalves (2009) afirma que os Sistemas Especialistas são uma das formas do ramo da IA que se baseia na especialização humana, no qual esses sistemas podem possuir domínio e experiência específica em alguma determinada área do conhecimento, sendo essa área medicina, logística, administração, dentre outras. Esses sistemas projetam as soluções e as tomadas de decisões com base no que foi passado a eles como conhecimento, agindo assim, na busca de soluções de problemas e sendo capazes de oferecer sugestões e conselhos aos usuários.

Para Wangenheim (2003), o Raciocínio Baseado em Casos (RBC) é uma abordagem para a solução de problemas e para o aprendizado com base em experiência passada. De uma forma simplificada, pode-se compreender o RBC como a solução de novos problemas por meio da utilização de casos anteriores já conhecidos. Um novo problema que é apresentado a essa forma de aprendizado é resolvido com a reutilização da solução de um problema anterior parecido com o atual, podendo ser uma solução aplicada totalmente ou parcialmente no novo problema.

Segundo Rosa e Luz (2009), Algoritmos Genéticos são técnicas que se baseiam na teoria de Darwin sobre a evolução genética dos seres vivos que tem como base a adaptação e o comportamento dos seres vivos. Com esse conceito, separa-se uma população que pode ser usada como uma possível solução de um problema, essa população é analisada e submetida a testes verificando sua adaptabilidade, depois essa população sofre um processo de *crossover* para gerar populações mais adaptadas, esse processo se repete até se encontrar uma solução satisfatória para o problema.

A teoria dos conjuntos *fuzzy* é uma extensão da teoria clássica de conjuntos (que trabalha apenas com grau de pertinência binária - 0 ou 1, verdadeiro ou falso). Essa teoria atribui graus intermediários de pertinência, entre 0 e 1, a cada um dos elementos dos seus conjuntos. Os modelos de inferência *fuzzy* são capazes de processar informações imprecisas e qualitativas e são adequados em processos que exigem tomadas de decisão. Para a representação *fuzzy* do conhecimento impreciso utilizam-se variáveis linguísticas, que admitem como valores expressões linguísticas, também denominadas de termos primários, como “muito grande”, “ruim”, “muito ruim”. Os termos primários de cada variável linguística formam a partição fuzzy destas variáveis (GOMIDE e GUDWIN, 1994).

A base de uma Rede Neural Artificial é imitar as ligações dos neurônios que existem no cérebro humano. As redes neurais são formadas de acordo com o conceito do cérebro humano, sendo elas desenvolvidas por meio de diversas ligações de neurônios (por isso a

ideia de rede) na qual o conhecimento é transmitido por intermédio da conexão entre eles. Na próxima seção será apresentado o conceito.

1.1.2 Redes Neurais Artificiais

Segundo Fernandes (2010), as Redes Neurais Artificiais procuram simular o funcionamento do cérebro humano imitando a constituição de neurônios que são compostos de massa central, e de dois tipos de filamentos, os axônios, transmissores de eletricidade, e os dendritos, receptores de impulsos elétricos. No encontro desses filamentos existe um espaço onde se processam reações químicas, conhecidas como sinapses. Essas sinapses são as responsáveis pelo aprendizado. O estudo neurológico auxilia no desenvolvimento dos neurônios artificiais para a sua utilização em IA.

Assim, a base principal das redes neurais é a simulação de neurônios do cérebro humano para neurônios sistêmicos, que se comunicam por meio de ligação (por isso o conceito de “rede”) (ZANELATO et al, 2017). Para que a rede neural funcione é necessária a utilização de algoritmos que aprendam com base em exemplos passados. Conforme a rede neural recebe as informações necessárias, é possível desenvolver máquinas mais inteligentes (MARSLAND, 2014).

A Figura 1 representa um neurônio artificial. Existem os vetores de entradas (X_n) e seus respectivos pesos numéricos (W_n), que representam as sinapses do cérebro; a função somadora que é calculada pela soma ponderada das entradas e submetida à uma função de transferência, ou função de ativação, que determina se a soma é maior que um valor numérico - o limiar do neurônio - se sim, o neurônio é ativado (valor 1) caso contrário, continuará desativado (valor 0). Pode também ser utilizada um valor limiar denominado vias (ou *bias*) que são ajustáveis. Para minimizar o erro entre a saída da rede neural e a saída esperada por padrão de treinamento, os pesos e as *bias* são ajustados por uma função.

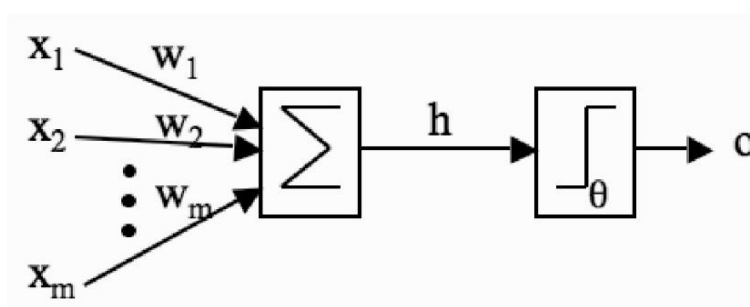


Figura 1: Modelo do neurônio artificial proposto por McCulloch e Walter Pitts

Fonte: Marsland, 2014. Pág: 41

Inspirado no modelo do neurônio artificial de McCulloch e Pitts, Frank Rosenblatt desenvolveu o primeiro modelo de uma arquitetura de rede neural artificial, o Perceptron.

O Perceptron é o modelo mais simples de uma arquitetura de uma rede neural artificial e o que deu embasamento para as demais redes neurais artificiais, sendo representado matematicamente, o Perceptron recebe inúmeras entradas e gera uma única saída.

Segundo Géron (2017) cada neurônio é caracterizado pelo peso, bias e a função de ativação. Os dados de entrada são alimentados na camada de entrada. Os neurônios fazem uma transformação linear na entrada pelos pesos e bias. A transformação não linear é feita pela função de ativação. A informação se move da camada de entrada para as camadas ocultas. As camadas ocultas fazem o processamento e enviam a saída final para a camada de saída. Este é o movimento direto da informação conhecido como propagação direta.

Quando os resultados esperados pela rede neural não forem aceitáveis (não são o que se esperava), então seriam atualizados os pesos e bias dos neurônios com base no erro. Este processo é conhecido como backpropagation. Uma vez que todos os dados passaram por este processo, os pesos e bias finais são usados para previsões (GÉRON, 2017).

Na rede neural, os pesos são considerados como a variável da função, é algo que sempre será calculado para adquirir os resultados e a cada interação será alterado, obtendo novos pesos para que o aprendizado se torne mais efetivo e assertivo. Já o modelo é a função em si, a própria rede neural, ou seja, quando se escreve a rede neural automaticamente está sendo escrito o modelo que será usado para o treinamento e esse modelo é salvo para não ser necessário ser reescrito o tempo todo.

Na Figura 2, pode-se visualizar como funciona a arquitetura de uma rede neural, na qual são inseridas as “entradas” que se deseja para o aprendizado, essas entradas são divididas em camadas gerando assim os pesos necessários para a interpretação do que se está recebendo e após esse processo, é gerada a “saída” esperada que já foi pré-definida anterior.

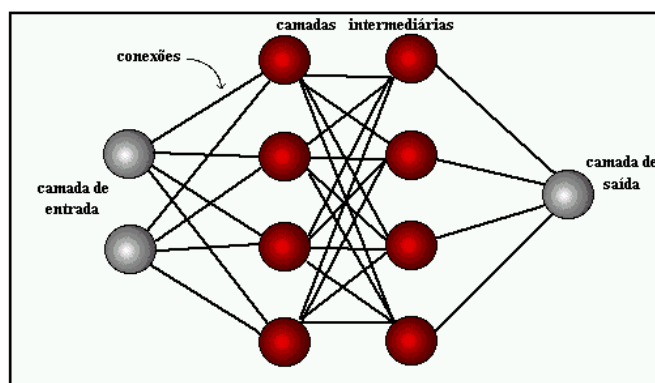


Figura 2: Arquitetura de uma rede neural

Fonte: <http://conteudo.icmc.usp.br/pessoas/andre/research/neural/>

Toda rede neural artificial precisa da utilização de função de ativação e outras funções que auxiliam na formação dos pesos e dos erros para assim gerar o aprendizado da rede neural artificial.

Segundo Palmiere (2016), a rede neural artificial contém cinco mecanismos de aprendizado (que utilizam as funções citadas anteriormente), sendo eles: supervisionado, não supervisionado, com reforço, lote padrão e padrão por padrão.

- a. **Supervisionado:** Esse tipo de aprendizado baseia-se em saber as amostras de entrada e saber como será exatamente as suas saídas, pois tendo esse conhecimento, os pesos e os limiares serão ajustados o tempo todo pelo algoritmo de aprendizado da rede neural artificial.
- b. **Não supervisionado:** Esse tipo de aprendizado baseia-se em saber quais são as amostras de entrada, porém não se conhece as saídas desses resultados, para isso é organizado os padrões e criado subconjuntos de acordo com as particularidades das amostras. Nesse aprendizado os pesos e os limiares são ajustados pelo algoritmo da rede neural artificial para refletir essas particularidades.
- c. **Com Reforço:** Semelhante a forma de treinamento supervisionado, propõe-se o algoritmo de ajuste aos pesos e os limiares de acordo com as informações adquiridas pelo sistema que mapeiam as determinadas entradas.
- d. **Lote de Padrões (Off-Line):** Primeiramente é gerado todo o conjunto de amostra de entrada para depois regularizar os pesos e os limiares.
- e. **Padrão por Padrão (On-Line):** Assim que a amostra de entrada é apresentada ao algoritmo, o mesmo reajusta os pesos e os limiares e após esse ajuste, essa amostra é descartada. Esse tipo de aprendizado é usado quando as amostras sofrem variações ao longo do tempo.

1.1.2.1 Medidas de Desempenho dos Classificadores

Para medir o desempenho dos modelos de classificação, pode ser utilizada uma matriz de confusão, que resume as informações sobre as classes corretas e aquelas preditas por um sistema de classificação. Conforme apresentado na Tabela 1, os resultados são apresentados em duas dimensões: classes verdadeiras e classes preditas, para k classes distintas $\{C_1, C_2, \dots, C_k\}$. Cada elemento $M(C_i, C_j)$ da matriz, $i, j = 1, 2, \dots, K$, representa o número de exemplos da base de dados que pertencem à classe C_i , mas que foram classificados como sendo da classe C_j .

Tabela 1 - Matriz de Confusão de um Classificador - para K classes

Classes	Predita C1	Predita C2	...	Predita Ck
Verdadeira C1	M (C1, C1)	M (C1, C2)	...	M (C1, Ck)
Verdadeira C2	M (C2, C1)	M (C2, C2)	...	M (C2, Ck)
...
Verdadeira Ck	M (Ck, C1)	M (Ck, C2)	...	M (Ck, Ck)

Fonte: MONARD e BARANAUSKAS, 2003, p. 102

O número de acertos, para cada classe, está localizado na diagonal principal da matriz, M (Ci, Ci). Os demais elementos M (Ci, Cj), em que $i \neq j$, representam erros na classificação.

Várias medidas podem ser derivadas da matriz de confusão. Para exemplificar essas medidas, apresenta-se, na Tabela 2, por simplicidade, uma matriz para um problema com apenas duas classes, denominadas como C+ (positiva) e C- (negativa). Nesse caso existem duas possibilidades de acerto: Verdadeiro Positivo (VP) e Verdadeiro Negativo (VN) e duas possibilidades de erro: Falso Positivo (FP) e Falso Negativo (FN).

Tabela 2 - Matriz De Confusão - para 2 classes

Classes	Predita C+	Predita C-
Verdadeira C+	VP	FN
Verdadeira C-	FP	VN

Fonte: Adaptado de Goldschmidt e Passos, 2005

Assim, as entradas da matriz de confusão têm os seguintes significados:

- VP é o número de predições corretas da classe C+;
- FN é o número de predições incorretas da classe C+;
- FP é o número de predições incorretas da classe C-;
- VN é o número de predições corretas da classe C-.

As equações de 1 a 6 são medidas obtidas a partir da matriz de confusão:

- Acurácia (AC) é a proporção do número total de predições que foram corretas:

$$AC = (VP + VN) / (VP + VN + FP + FN) \quad (1)$$

- Sensitividade, Revocação, ou Taxa de Verdadeiro Positivo (TVP) é a proporção de casos positivos que foram identificados corretamente:

$$TVP = VP / (VP + FN) \quad (2)$$

- Taxa de Falso Positivo (TFP) é proporção de casos negativos que foram classificados incorretamente como positivos:

$$TFP = FP / (FP + VN) \quad (3)$$

- Especificidade ou Taxa de Verdadeiro Negativo (TVN) é definida como a proporção de casos negativos que foram classificados corretamente:

$$TVN = VN / (VN + FP) \quad (4)$$

- Taxa de Falso Negativo (TFN) é a proporção de casos positivos que foram classificados incorretamente como negativos:

$$TFN = FN / (FN + VP) \quad (5)$$

- Precisão (P) é a proporção de casos positivos preditos que foram corretos:

$$P = VP / (VP + FP) \quad (6)$$

De acordo com Han e Kamber (2006) as taxas de verdadeiro positivo, verdadeiro negativo, falso positivo e falso negativo são úteis para avaliar os custos (ou riscos) e benefícios associados com o modelo de classificação.

Após o conceito sobre a forma que uma rede neural artificial realiza seu treinamento para gerar o conhecimento necessário para solucionar um determinado problema, é necessário a inserção de *deep learning* para aprimorar o conhecimento das redes neurais mais complexas. Esses conceitos são apresentados a seguir.

1.1.2.2 Deep Learning

De acordo com Goodfellow, Bengio e Courville (2016) *deep learning* (aprendizado profundo) é um ramo da Inteligência Artificial (IA), utilizado para soluções de problemas por meio de aprendizado e entendimento de uma hierarquia de conceitos simples, por exemplo: o céu é azul. Esse é um conceito simples, mas quando é inserido na *deep learning*, ao se adentrar no conceito mais complexo e obter o seguinte questionamento, por exemplo: por que o céu é azul? *Deep learning* auxilia na procura da resposta para tal pergunta, pois busca compreender mais a fundo um determinado assunto.

Assim, os métodos que usam *deep learning* procuram interpretar e descobrir um modelo. Para isso, usa-se um conjunto de dados (que são as entradas) e um método para conduzir o aprendizado do modelo a partir desses dados. Quando o processo de aprendizado é finalizado, é possível obter uma função que é capaz de receber os dados brutos (como forma de entrada, parâmetros iniciais) e fornecer como saída uma representação apropriada para o problema em questão (GOODFELLOW, BENGIO e COURVILLE 2016).

As redes neurais artificiais mais complexas utilizam *deep learning* para tornar o aprendizado e o conhecimento dessas redes mais precisos e com soluções de problemas mais complexos.

1.1.2.3 Redes Neurais Convolucionais

As redes neurais convolucionais são comumente utilizadas na classificação de imagens e na detecção de objetos. (GÉRON, 2017).

Uma Rede Neural Convolucional é um algoritmo de Aprendizado Profundo que pode captar uma imagem de entrada, atribuir importância (pesos e bias que podem ser aprendidos) a vários aspectos / objetos da imagem e ser capaz de diferenciar um do outro. (GÉRON, 2017).

O pré-processamento exigido em uma Rede Neural Convolucional é muito menor em comparação com outros algoritmos de classificação. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, a Rede Neural Convolucional têm a capacidade de aprender esses filtros / características. (GÉRON, 2017).

A arquitetura de uma Rede Neural Convolucional é análoga àquela do padrão de conectividade de neurônios no cérebro humano e foi inspirada na organização do Visual Cortex. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida como Campo Receptivo. (GÉRON, 2017).

O modelo de uma Rede Neural Convolucional está representado na Figura 3 a seguir.

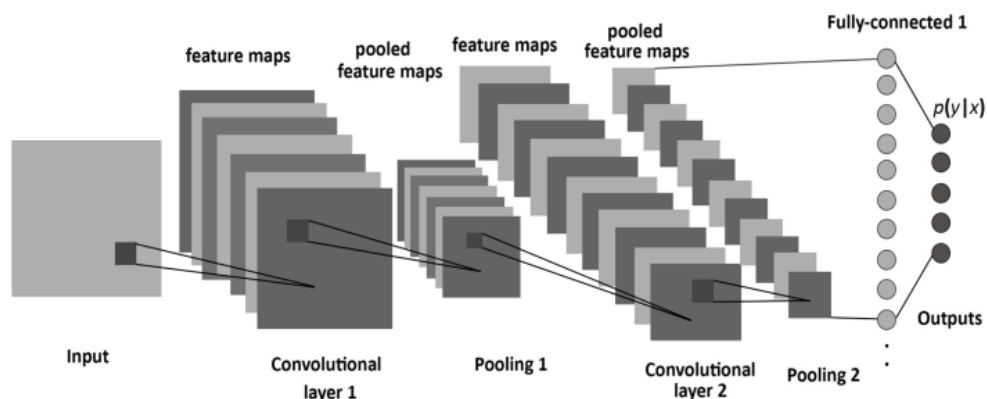


Figura 3: Típico diagrama em bloco de uma Rede Neural Convolucional

Fonte: <https://appliedmachinelearning.blog/2018/03/24/achieving-90-accuracy-in-object-recognition-task-on-cifar-10-dataset-with-keras-convolutional-neural-networks/>.

De acordo com Kumar (2018), a camada de convolução é composta por 3 operações: campos receptivos locais, pesos compartilhados e *pooling*.

Na camada de campos receptivos locais cada pixel de uma imagem é um valor numérico que representa a intensidade de cor de acordo com a escala de cor utilizada, como

RGB (Red, Green, Blue), por exemplo, ou apenas intensidade em escala de cinza para imagens em preto e branco.

Os pesos compartilhados são auxiliados para realizar os cálculos com base nos resultados das outras camadas.

De acordo com Géron (2017) na Rede Neural Convolucional existem camadas de Pooling, que são geralmente usadas imediatamente após camadas convolucionais e o que fazem é simplificar as informações na saída da camada convolucional, ou seja, uma camada de pooling que recebe cada saída do mapa de características da camada e prepara um mapa de características condensadas.

Como uma rede neural convolucional é a mais indicada na tratativa de reconhecimento facial, por trabalhar muito bem com a identificação de imagens.

A seguir será apresentado como o reconhecimento facial pode ser feito por meio de pontos estratégicos na face.

1.1.3 Reconhecimento facial

O reconhecimento facial acontece naturalmente nos seres humanos, pois os olhos captam as imagens das faces e o cérebro as identificam, assim, para um ser humano é algo relativamente simples reconhecer a emoção que outro indivíduo está sentindo pela sua expressão facial. As emoções mais comuns de se reconhecer são: alegria, tristeza, neutra, surpresa, medo e raiva (BUSSO et al., 2004; KELTNER et al., 2003 apud ZANELATO et al., 2017).

Ekman e Friesen (1978, apud Dantas et al, 2015), descreveram alguns dos possíveis movimentos faciais visíveis conforme o indivíduo mexe os músculos da face. Cada emoção exerce uma contração em algum músculo (a boca, os olhos, o nariz, sobrancelhas), gerando assim uma feição diferente. Para identificar cada movimento que a face realiza de acordo com as emoções, foram utilizados 62 Unidades de Ação (UA) e cada UA tem seu código numérico associado a um movimento da face.

Os devidos pontos estratégicos são marcados nos rostos das pessoas das imagens, facilitando assim a identificação da expressão de cada pessoa, com base na movimentação que os pontos fazem nas regiões estratégicas (sobrancelhas, boca, olhos e nariz).

Na Figura 4, é possível visualizar um exemplo de como é realizada a marcação dos pontos estratégicos na face de um ser humano por meio de bibliotecas que tratam imagens e vídeos.



Figura 4: Pontos na face

Fonte: Techtudo (<https://www.techtudo.com.br/artigos/noticia/2012/04/como-funciona-o-reconhecimento-facial.html>)

Costa (2017, p.31, apud Teixeira 2017) afirma que cada UA é a contração ou relaxamento de um músculo ou uma combinação de músculos da face. Por exemplo, a UAc1 represente uma sobrancelha levantada e a UAc2 uma sobrancelha fechada e contraída. Assim, as UAs são independentes de interpretação e podem ser usadas, por exemplo, para o reconhecimento de emoções básicas como alegria, tristeza, raiva, surpresa, medo e neutra.

Segundo Taigman et al (2014) o reconhecimento facial pode ser dividido em quatro partes, sendo elas: detecção, alinhamento, descrição e classificação:

- a. **Detecção:** É a utilização de um algoritmo para identificar faces em uma imagem, basicamente se passa um filtro pela imagem e utiliza a classificadores para determinar se a imagem apresentada representa uma face ou não.
- b. **Alinhamento:** Consiste em estimar a orientação da face e a posição dos olhos, da boca, das sobrancelhas e do nariz.
- c. **Descrição:** Consiste no desenvolvimento um descritor para a face humana de tal maneira que ao utilizar um classificador (próximo passo), possa ter uma acurácia alta.
- d. **Classificação:** Essa etapa existe um aprendizado de máquina profundo para encontrar a melhor forma de descrever essas faces.

Os algoritmos de reconhecimento facial utilizam esses diversos pontos da face para avaliar a expressão. Ao treinar uma Rede Neural com imagens com várias expressões, segundo Zanellato, et al (2017), a comparação das expressões pode ser feita por meio de *frames*, que com o auxílio de um fluxo óptico possibilita a análise dos movimentos dos músculos do rosto, comparando o tempo decorrido do movimento desse músculo e os

movimentos da cabeça, permitindo que certas áreas consideradas mais sutis para serem notadas, possam ser analisadas mais facilmente.

Analisando a forma como o reconhecimento facial é realizado, pelos pontos estratégicos, precisa-se entender como funciona as emoções do ser humano.

1.1.4 Inteligência emocional

Para compreender como a inteligência emocional funciona, é necessário primeiramente, entender o que é a emoção em si.

As emoções sempre desempenharam um papel importante na evolução e na adaptação do ser humano no meio ambiente em geral. Essas emoções também iluminam o juízo das pessoas alertando de riscos e consequentemente, acionando o sistema de tomada de decisões. A cada segundo, o cérebro recebe bilhões de informações com relação aos sinais que recebemos do dia a dia e essas informações influenciam outros fenômenos psicológicos, tais como atenção, memória, linguagem verbal e não verbal.

De acordo com Lino (2004), quando adentra-se no assunto inteligência emocional, compreende-se que está relacionada à parte emocional como um todo e de forma abrangente, pois se encontra em todos os momentos de vidas, seja na forma de convivência familiar, social e de trabalho. O ser humano tem a sua primeira oportunidade de estar dentro de um ambiente empresarial, sente-se confuso e perdido, começa a se questionar de todas as responsabilidades que vai passar a ter a partir daquele momento, diante disso, desenvolve o sentimento de confusão por ser algo novo e começa a aprimorar seus sentidos para a responsabilidade. Mediante o exposto, percebe-se que no decorrer dessa evolução o indivíduo começa a procurar se aperfeiçoar para melhorar o desempenho dentro do seu setor de trabalho. Quando o indivíduo se encontra balanceado e mais seguro de si, ele amplia a sua visão e seus sentimentos são mais controlados, gerando assim uma eficiência em suas ações.

Segundo Woyciekoski e Hutz (2009) a emoção corresponde a uma reação psicobiológica complexa, que envolve inteligência e motivação, impulso para ação, além de aspectos sociais e da personalidade, que acompanhados de mudanças fisiológicas, expressariam um acontecimento significativo para o bem-estar subjetivo do sujeito no seu encontro com o ambiente. Por meio das emoções que o ser humano sente, é possível gerar funções comunicativas e sociais para que haja uma interação maior entre as pessoas. Por meio do desenvolvimento dessa inteligência as pessoas se tornam mais prudentes, intuitivas e mais equilibradas para tomada de decisões mais assertivas.

A inteligência emocional se torna uma condição indispensável para o planejamento de atividades no meio empresarial no que se refere à gestão e organização, sendo visível que é por meio dela que haverá a possibilidade dos profissionais contribuírem para o desenvolvimento da empresa. Esses profissionais que praticam da inteligência emocional se tornam melhores em trabalho em equipe, pois compreendem a si mesmos e ao próximo, tendo ainda mais empatia, bem como motivando-se constantemente no sentido de alcançar resultados extraordinários para os negócios como um todo.

A junção do reconhecimento das emoções faciais e a compreensão sobre inteligência emocional podem ajudar empresas a conseguir identificar algum tipo de mudança repentina de humor com um respectivo colaborador e assim notificar um superior sobre essa alteração, auxiliando o superior a alguma tomada de decisão ou até mesmo uma conversa para identificar um possível problema que ocorreu com o colaborador.

1.2 Trabalhos relacionados

As pesquisas têm por objetivo exemplificar a forma que outros trabalhos realizaram o aprendizado de máquina por meio de Inteligência Artificial (IA), algoritmos e até mesmo metodologias, focando no reconhecimento de emoções faciais mediante a utilização de imagens e vídeos.

Diniz et al (2013) desenvolveu um sistema para realizar o reconhecimento facial dos seres humanos. Para que isso fosse possível, o autor dividiu seu desenvolvimento em duas etapas, a primeira utilizando PCA (Principal Components Analysis - Análise dos Componentes Principais) e Eigenfaces (autofaces) para extrair algumas características da face; e a segunda utilizando classificadores K-Nearest Neighbors (K-NN), Random Forest (Floresta Aleatória) e K-Star (K-estrela) para reconhecimento da face. Para testar esse algoritmo, foi utilizada uma base de dados com 1280 imagens de 64 classes diferentes. Os resultados obtidos no experimento foram taxas acima de 90% em todos os classificadores.

A pesquisa de Dantas et al (2015), levanta o resultado do experimento que detectou sete emoções básicas do ser humano (alegria, tristeza, neutra, surpresa, desgosto, raiva e medo), capturadas por meio de pontos de referências e movimentação facial que expressam algum tipo de reação em imagens e vídeos. Para o reconhecimento das emoções foi utilizada a técnica de Árvore de Decisão, além da utilização de algoritmos que validavam as emoções em tempo real. Os autores afirmaram que os resultados da pesquisa foram satisfatórios para o reconhecimento de todas as emoções mencionadas acima, atingido um percentual geral de 86,4% de reconhecimento.

A pesquisa de Guimarães (2015) objetivou validar a eficácia e viabilidade do uso do Kinect¹ para reconhecimento biométrico de características faciais no ambiente empresarial, além de desenvolver um protótipo para o reconhecimento dessas características em tempo real para o registro e controle do ponto em empresas. Para o desenvolvimento do software os autores utilizaram: a linguagem de desenvolvimento CSharp, a ferramenta de compilação Visual Studio 2010 (ambos da Microsoft), o gerenciador de banco de dados MySQL e uma biblioteca OpenCV (*Open Source Computer Vision Library*), além da parte de software foi necessária a utilização do *Kinect* com um webcam de alta definição para captação das características da face do ser humano, gerando imagens para análise. Após vários testes com o protótipo desenvolvido, os autores afirmaram que a utilização do *Kinect* foi essencial para a captura das imagens no momento em que indivíduo entrava no recinto.

O projeto de Sousa, et al (2016) apresentou os resultados obtidos na avaliação dos aspectos afetivos do computador, ou seja, qual a reação de um ser humano ao utilizar certos aplicativos de um computador. O projeto foi feito com 4 emoções faciais (alegria, raiva, surpresa e tristeza). Os traços da personalidade foram capturados pelo Kinect para o desenvolvimento de uma ferramenta capaz de identificar a satisfação do usuário, sendo essa satisfação positiva ou negativa. Para isso, os autores desenvolveram uma ferramenta que consiste em: Captura de Imagens pelo Kinect, Rastreamento das Características faciais, Classificação e Inferência das emoções da face. As afirmações dos autores descreveram que o projeto desenvolvido por eles pode ser usado para reconhecer a face em tempo real na utilização de novos sistemas, para reconhecer a aceitação da publicação.

A pesquisa de Silva (2016) buscou demonstrar como a utilização de classificadores supervisionados Floresta de Caminhos Ótimos (OPF – uma família de classificadores para auxiliar na precisão do reconhecimento facial) podem ser assertivos na comparação de duas imagens de uma face humana para reconhecer se ambas pertencem ao mesmo indivíduo. O autor utilizou duas imagens de indivíduos e os classificou em dois grupos: a classe que representa imagens diferença do mesmo indivíduo (classe intrapessoal) e a classe que representa imagens diferença de indivíduos diferentes (classe interpessoal). A partir disso, houve uma alteração no método de decisão do OPF para adaptá-lo e transformá-lo em uma função que pudesse definir a semelhança entre cada classe selecionada, além de utilizar os métodos Eigenfaces e EBGM para validar se duas imagens distintas são de uma mesma

¹ O Microsoft Kinect (MICROSOFT, 2011) é um sensor de movimento lançado em novembro de 2010 capaz de fornecer várias informações em tempo real, como a distância de uma pessoa do dispositivo, assim como uma imagem do seu esqueleto. A tecnologia de captura de dados em tempo real permite utilizar o dispositivo em sistemas biométricos que necessitem de respostas rápidas.

pessoa, para identificar se foi executado com sucesso. Obtendo resultados que classificadores podem ser modificados e auxiliados no uso de reconhecimento de pessoas.

Furtado et al (2016) propuseram a utilização de técnicas de redes neurais com deep learning para reconhecer os integrantes de um elenco de filmes, pelas suas características faciais em vídeos como trailers. Os autores utilizaram um algoritmo de alinhamento tridimensional para reconhecimento das faces de 2606 celebridades por meio de imagens de suas fisionomias, foi utilizado para treinar uma Rede Neural Convulacional (RNC) visando a aprender e identificar melhor as faces dos indivíduos com base nas imagens adquiridas. Após treinar essa RNC, foi utilizado para reconhecer e identificar atores em trailers de filmes. De acordo com os autores os resultados foram promissores, podendo ser uma solução viável para aprofundar o conhecimento e o uso de deep learning para reconhecimento facial em vídeos.

O estudo de Teixeira (2017) teve por objetivo reconhecer quatro emoções básicas (alegria, tristeza, surpresa e raiva) do ser humano por meio da análise facial. Foram utilizadas técnicas de *clusterização* (Expectation Maximization (EM) e K-means) e uma base de dados com valores de Unidades de Animação (UnA) gratuita disponível para download. Os dados os, foram analisados por meio da ferramenta WeKa com algoritmos de classificação e associação. A ferramenta Fourface foi usada para captura das expressões faciais dos usuários para validar a viabilidade da mineração de dados com algoritmos para reconhecer as emoções. Após avaliação dos resultados das capturas das expressões e as ferramentas utilizadas, o autor considerou um acerto médio de 67,9% das expressões selecionadas para teste.

A pesquisa de Zanellato et al (2017) pretendeu identificar, por meio de fotos da face frontal do ser humano, sete expressões básicas (felicidade, tristeza, medo, raiva, surpresa, aversão e neutro). Para que esse experimento fosse possível, os autores escolheram o treinamento de deep learning de duas camadas: a Rede Neural Convulacional, capaz de extrair características de imagens e a Rede Interconectada para a classificação dessas expressões. Com base nos algoritmos classificadores, foi realizado o aprendizado da máquina para reconhecer as faces em tempo real, para que assim, o sistema desenvolvido pelos autores fosse eficaz. Por meio desses experimentos, os resultados surpreenderam as expectativas dos autores, por chegarem acima de 92,24% em média esperada.

O trabalho de Dantas et al (2018), teve por objetivo desenvolver um aplicativo móvel para pessoas com deficiência visual para auxiliar essas pessoas a detectar 6 sentimentos básicos do ser humano (alegria, tristeza, surpresa, desgosto, raiva e medo). Para auxiliá-lo nesse desenvolvimento, os autores utilizaram algoritmos Haar Cascade (que utiliza cálculo da integral de imagem, possuindo os melhores tempos de resposta para essas aplicações) que é

otimizado pelo classificador chamado Adaboost (otimiza o tempo de resposta e economiza recursos computacionais), foi utilizado imagens em tons de cinza, para minimizar o ruído que as imagens causavam. O dispositivo vibra de forma diferente para que pessoas com deficiência visual possam reconhecer as expressões em um ambiente social. Os autores afirmaram que 88% das expressões foram reconhecidas pelo dispositivo validando-o para ser usado com pessoas com deficiência visual.

O experimento de Damasceno et al (2018) avaliou o conhecimento das pessoas sobre o conceito e a percepção do que é Inteligência Artificial (IA) e sobre seus impactos em um futuro próximo. Para isso esse experimento, foi selecionado 18 pessoas entre 20 até 30 anos de idades, essas pessoas foram usadas para identificar o conhecimento sobre o conceito de IA. A pesquisa tinha quatro questões objetivas e uma questão subjetiva. Com os resultados obtidos pela pesquisa, os autores perceberam que é necessário melhorar o conhecimento popular sobre IA, pelo baixo conhecimento que a população tinha sobre esse assunto. O experimento de Damasceno, et al auxiliou na identificação do conhecimento popular sobre IA verificando a necessidade de ampliar esse conhecimento, não deixando fazer parte apenas de conceitos de ficção científica.

As técnicas utilizadas nos trabalhos relacionados, auxiliam no embasamento da ideia deste projeto, trazendo soluções utilizadas por outros que possam ser aperfeiçoadas e melhoradas para atingirem resultados mais satisfatórios do que os atuais na área de Inteligência Artificial. Com base nesses trabalhos, percebe-se que os que mais se aproximam da ideia para introduzir o reconhecimento facial por meio de técnicas semelhantes ao utilizada nesse trabalho, são Diniz (2013), Dantas (2015), Teixeira (2017), Dantas (2018). Em seguida apresenta-se, resumidamente, no Quadro 2 as informações referentes trabalhos relacionados.

Quadro 2: Trabalhos Relacionados

Autor/Ano	Objetivo	Ferramentas	Algoritmos
Diniz, Neto, Júnior, Fontes (2013)	Reconhecer a face utilizando classificadores.	Weka	PCA, Eigenfaces K-Nearest Neighbors (K-NN), Random Forest e K-Star
Dantas, Melo, Moura, Fernandes (2015)	Realizar o reconhecimento dinâmico de emoções	-	Árvore de Decisões e <i>Haar Cascade</i>
Guimarães (2015)	Desenvolver um protótipo de software de reconhecimento facial de tempo real	<i>Kinect, Webcam</i>	CSharp, Open CV, MySQL
Souza, Costa, Pires, Araújo (2016)	Reconhecer as emoções faciais enquanto usa aplicativos computacionais	<i>Kinect, Webcam</i>	-
Silva (2016)	Combinar classificadores para reconhecimento facial	-	Eigenfaces, EBGM e Famílias de classificadores
Furtado, Souza (2016)	Descobrir elenco em trailers com deep learning	-	Rede Neural Convulacional
Teixeira (2017)	Reconhecer a face por meio de mineração de dados	Weka	Expectation Maximization (EM) e K-means
Zanellato, Ferreira (2017)	Desenvolver um sistema de reconhecimento de expressões faciais	-	Rede Neural Convulacional e Rede Interconectada
Dantas, Gouveia, Melo, Gonçalves, Nascimento (2018)	Desenvolver um aplicativo móvel para deficientes visuais reconhecerem emoções	Emotion Visual	<i>Haar Cascade</i> e <i>Adaboost</i>
Damaceno, Vasconcelos (2018)	Identificar o conhecimento popular sobre inteligência artificial	-	-

Fonte: Elaborado pelo autor.

CAPÍTULO II

Metodologia

2.1 Natureza da Pesquisa

No desenvolvimento desse estudo, foi utilizado a metodologia de Pesquisa Experimental a qual, de acordo com Gil (2002), consiste na determinação de um objeto de estudo, selecionar as variáveis capazes de influenciá-lo e definir meios de controle e observação dos efeitos que esta variável produz neste objeto.

2.2 Variáveis de análise

No planejamento e implementação deste trabalho, foram referenciadas duas variáveis:

1. Qualidade do banco de imagem.
2. Parâmetros de treinamento da rede neural.
3. Acurácia da rede neural.
4. Eficácia de registro das emoções.

2.3 Ferramentas Utilizadas

As tecnologias e ferramentas utilizadas para o desenvolvimento deste trabalho são descritas a seguir:

- Python²: é uma linguagem de programação que pode ser utilizada para diversos aspectos, mas nessa pesquisa, essa linguagem será usada para desenvolver a rede

² Conforme disponível em: <<https://www.python.org/doc/>>. Acesso em: 20 ago. 2019.

neural, com base nos estudos, percebeu-se que essa linguagem é mais prática e fácil de se utilizar em projetos que implementam inteligência artificial. O desenvolvimento e adaptação da rede neural foi realizada com Python.

- Anaconda Navigator³: é uma ferramenta que auxilia no desenvolvimento de aplicações de aprendizado de máquina, instalando o navegador e usando o prompt de comando dele mesmo para executar importação de bibliotecas e a utilização do aprendizado, usando recursos (processamento e memória) da própria máquina.
- Open Source Computer Vision (OpenCV)⁴: é uma biblioteca utilizada para manipulação de imagens e vídeos.
- Keras⁵: é uma API de rede neural de alto nível, escrita em Python. Foi desenvolvida para auxiliar no aprendizado de máquina. Permite prototipagem fácil e rápida (através da facilidade de uso, modularidade e extensibilidade). Suporta redes convolucionais e redes recorrentes, bem como a combinações das duas.

2.4 Experimento de Pesquisa

O experimento desta consistiu na aplicação de um algoritmo de reconhecimento de expressões e emoções faciais treinados a partir de Redes Neurais. Foram utilizadas as 7 emoções básicas raiva, nojo, medo, feliz, triste, surpresa e neutro.

Utilizou-se uma Rede Neural Convolucional supervisionada que passou por um processo de aprendizagem/treinamento entradas e saídas definidas, esse processo de aprendizado foi explicado na seção 1.1.2.

Para iniciar o treinamento da rede neural foi definido um modelo e os pesos para a identificação de cada rosto e expressão contemplada na base de dados escolhida para o treinamento/aprendizagem. Para a identificação dos rostos e das expressões, o algoritmo utilizou os pontos na face das imagens da base de dados que tem as expressões desejadas, são pontos estratégicos (conhecidos como Unidades de Movimento – UM) colocados nos rostos dos indivíduos, esses pontos são colocados em locais que há mudança das articulações do rosto, por exemplo boca, nariz, olhos e sobrancelhas.

Cada ponto no rosto de um indivíduo contém um código numérico para identificar a ação em um rosto (essa ação é conhecida como Unidades de Ação – UA), para a marcação desses pontos são utilizadas bibliotecas que manipulam imagens e vídeos.

³ Conforme disponível em: <<https://www.anaconda.com/distribution/>>. Acesso em: 20 ago. 2019.

⁴ Conforme disponível em: <<https://opencv.org/>>. Acesso em: 20 ago. 2019.

⁵ Conforme disponível em: <<https://keras.io/>>. Acesso em: 20 ago. 2019.

A rede neural foi treinada utilizando uma base de dados adquirida pelo *Kaggle*⁶ com 35888 imagens de diversas pessoas com diferentes idades e gênero, com ordenação aleatórias, essa aleatoriedade visa ao aprimorando do aprendizado do modelo.

Quando o modelo e os pesos são salvos, é realizada a validação da aplicação. Nesse caso, essa rede neural foi usada para reconhecimento de faces em tempo real, por meio de uma webcam que existe nos computadores (a que é acoplada em notebooks), validando cada expressão feita pelo indivíduo.

Na Figura 5 é possível visualizar a arquitetura do treinamento e da validação da rede neural convolucional apresentada neste trabalho.

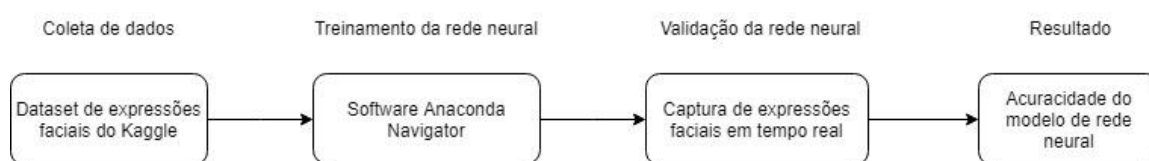


Figura 5: Arquitetura do projeto

Fonte: Elaborado pelo autor

2.5 Critérios para avaliação da pesquisa

A avaliação desta pesquisa foi realizada com três pessoas que trabalham em desenvolvimento de sistemas, que se voluntariaram para captura dos rostos em seu dia-a-dia no ambiente de trabalho. Com base nisso, foi verificada se as expressões de acordo com as 7 emoções básicas (raiva, nojo, medo, feliz, triste, surpresa e neutro) selecionadas para análise.

A partir da instalação e inicialização da aplicação na máquina das pessoas, suas faces eram capturadas pela *webcam* e com base na expressão daquele momento, a aplicação armazenava o resultado obtido. Esse processo era realizado no decorrer do dia, gravando-se 2 segundos de vídeo a cada 1 hora.

Próximo ao final do expediente de trabalho, o arquivo de log em que a aplicação armazenava seus resultados era utilizado pelo modelo gerado pela rede neural para verificar se seus resultados foram fieis ou não, de acordo com a acurácia gerada por essa verificação dos *frames* gerado pelos 2 segundos de vídeo.

A avaliação foi feita com três pessoas em seu ambiente de trabalho. Cada um trabalha 8 horas por dia, a validação foi realizada durante uma semana de trabalho que equivale a 5 dias. A cada 1 hora eram salvos 2 segundos de vídeo, equivalente a 50 *frames*. Multiplicando por 8 horas de trabalho equivale a 400 *frames*, vezes 5 dias por semana equivale a 2 mil fotos na semana, como o experimento foi realizado com 3 colaboradores, foram geradas 6 mil fotos para validar a rede neural.

⁶ Conforme disponível em: <<https://www.kaggle.com>>. Acesso em: 15 set. 2019.

CAPÍTULO III

Apresentação, implementação, testes e resultados

Neste capítulo, são apresentados os resultados obtidos para essa pesquisa, de acordo com a seguinte estrutura: a base de dados utilizada para realização do treinamento da rede neural, a forma que a pesquisa foi desenvolvida, utilizando a rede neural e os testes realizados, o uso da identificação em tempo real e os resultados apresentados.

3.1 Base de dados utilizada

Para contemplar o quesito de treinamento e teste da rede neural, foram utilizadas duas bases de dados distintas e ambas foram adquiridas gratuitamente pelo site do *Kaggle*. Esse site é focado no aprendizado de máquina (*machine learning*), existem inúmeros exemplos de *machine learning* e base de dados disponíveis para *download* gratuito de diversos estudos diferentes.

Foi necessário realizar um cadastro no site *Kaggle* com e-mail e senha ou a utilização de alguma rede social. Após realizar o registro no site, foi solicitado uma confirmação via e-mail, para efetivar o cadastro e permitir o acesso ao link de download dessa base de dados.

Para realizar o treinamento da rede neural, validando as expressões, utilizou-se a base de dados *Challenges in Representation Learning: Facial Expression Recognition Challenge*⁷.

A base de dados utilizada para o treinamento, concentra um total de 35.888 exemplos de imagens de rostos em escala de cinza de 48x48 pixels. Os rostos foram registrados automaticamente para que fique mais ou menos centralizado e ocupa-se aproximadamente a

⁷ Conforme disponível em: <<https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>>. Acesso em: 30 set. 2019.

mesma quantidade de espaço em cada imagem. A tarefa foi categorizar cada rosto com base na emoção mostrada na expressão facial em uma das sete categorias existentes na base de dados (0 = Raiva, 1 = Nojo, 2 = Medo, 3 = Feliz, 4 = Triste, 5 = Surpresa, 6 = Neutro).

O arquivo em CSV “Fer2013.csv” contém duas colunas, *emoção* e *pixels*. A coluna “emoção” contém um código numérico que varia de 0 a 6, esse código serve para a identificação da emoção que está sendo expressa pela imagem. A coluna “pixels” contém uma sequência entre aspas para cada imagem.

3.2 Processo de treinamento da rede neural convolucional

Para o desenvolvimento da pesquisa, foi utilizada uma aplicação pronta que foi extraída no repositório do Github⁸.

Após a definição da base de dados a ser utilizada nesse estudo, foi necessária a leitura do arquivo que continha as informações dos rostos com as respectivas expressões/emoções. Como explicado na seção 3.1, a base de dados foi primeiramente tratada para que o modelo da rede neural identificasse os rostos, transformando os pixels em imagens.

Depois da transformação dos pixels em imagens, o modelo da rede neural foi escrito utilizando a biblioteca do *Keras* e a linguagem de programação *Python* (explicadas na seção 2.4). A junção dessas duas ferramentas possibilitou a construção das camadas de identificação da rede neural, pois um modelo de uma rede neural é composto por um ou mais camadas, cada camada é escrita utilizando um aprofundamento maior com relação a cada época que a rede passa.

Para este estudo foi escrito quatro camadas do modelo da rede neural. As imagens passam por cada camada da rede e assim a rede compara os pontos, as semelhanças e analisa as saídas (que são as emoções de cada rosto). Quando o processo de treinamento é passado em cada camada da rede, os pesos são gerados por meio da função de ativação da rede neural, apresentando uma assertividade nos resultados da identificação das imagens e expressões. A quantidade de épocas que a rede neural irá tentar executar é definida no início no código do modelo, essa quantidade de épocas serve para definir quantas vezes essa rede neural passou pelas camadas para gerar o “conhecimento” da identificação.

⁸ Conforme disponível em:

<https://github.com/abhijeet3922/FaceEmotion_ID/blob/master/real_time_video.py>. Acesso em: 20 out. 2019.

Na Figura 6 é possível visualizar um trecho da rede neural passando em cada época. Observa-se ainda a porcentagem de acerto (*acc*) e perda (*loss*) sendo alterada no decorrer das épocas.

```
Epoch 00035: val_loss did not improve from 1.04294
Epoch 36/110
449/448 [=====] - 11s 26ms/step - loss: 1.0015 - acc: 0.6273 - val_loss: 1.1180 - val_acc: 0.5899

Epoch 00036: val_loss did not improve from 1.04294
Epoch 37/110
449/448 [=====] - 11s 25ms/step - loss: 0.9935 - acc: 0.6326 - val_loss: 1.0751 - val_acc: 0.6137

Epoch 00037: val_loss did not improve from 1.04294
Epoch 38/110
449/448 [=====] - 11s 25ms/step - loss: 0.9963 - acc: 0.6290 - val_loss: 1.0923 - val_acc: 0.6014

Epoch 00038: val_loss did not improve from 1.04294
Epoch 39/110
449/448 [=====] - 11s 25ms/step - loss: 0.9867 - acc: 0.6326 - val_loss: 1.0569 - val_acc: 0.6140

Epoch 00039: val_loss did not improve from 1.04294
Epoch 40/110
449/448 [=====] - 11s 25ms/step - loss: 0.9799 - acc: 0.6358 - val_loss: 1.0693 - val_acc: 0.6117

Epoch 00040: val_loss did not improve from 1.04294
Epoch 41/110
449/448 [=====] - 12s 26ms/step - loss: 0.9820 - acc: 0.6355 - val_loss: 1.0762 - val_acc: 0.6140
```

Figura 6: Treinamento da rede neural

Fonte: <https://appliedmachinelearning.blog/2018/11/28/demonstration-of-facial-emotion-recognition-on-real-time-video-using-cnn-python-keras/>

Quando esse processo foi finalizado, a rede neural apresentou uma acurácia em torno de 65%. O modelo e os pesos gerados por essa rede neural são salvos em um arquivo no formato de hdf5 (esse é o formato comum para armazenado de modelos e pesos de uma rede neural) por meio da chamada da função de *call-back*.

A partir do momento que o modelo e os pesos são definidos e salvos, é possível utilizá-los em outras aplicações.

Após o modelo e os pesos serem salvos, uma outra parte da aplicação foi adaptada utilizando a biblioteca do *OpenCV* e a linguagem *Python* (explicadas na seção 2.4). Essa aplicação faz a identificação em tempo real das expressões do indivíduo por meio da webcam, a biblioteca do *OpenCV* trabalha na tratativa de imagens e vídeos, por isso foi usada nesse processo da captação da imagem em tempo real.

3.3 Realização dos testes da aplicação

Para avaliar o modelo e os pesos gerados pela rede neural explicada na seção 3.2, foi desenvolvida uma aplicação com *OpenCV* para reconhecimento da face e da expressão em tempo real.

A aplicação valida as emoções, citadas na seção 3.1, sendo elas: raiva, nojo, medo, feliz, triste, surpresa e neutro.

Na Figura 7 é possível visualizar um teste feito em tempo real da rede neural, reconhecendo a expressão de raiva (angry).

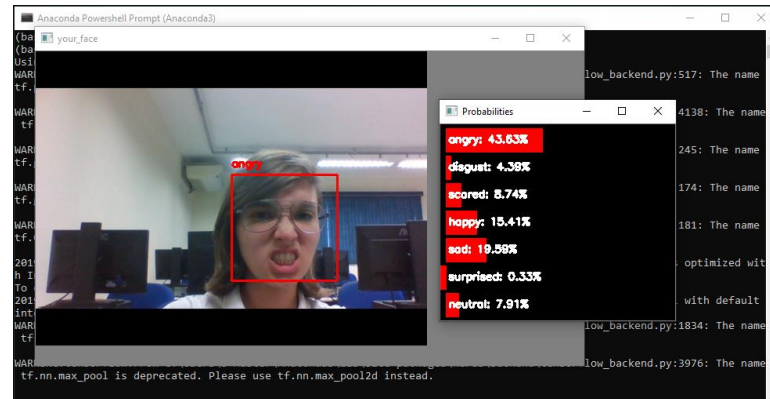


Figura 7: Reconhecimento da face e expressão: Raiva

Fonte: Elaborado pelo autor.

Na Figura 8 é possível visualizar um teste feito em tempo real da rede neural, reconhecendo a expressão de feliz (happy).

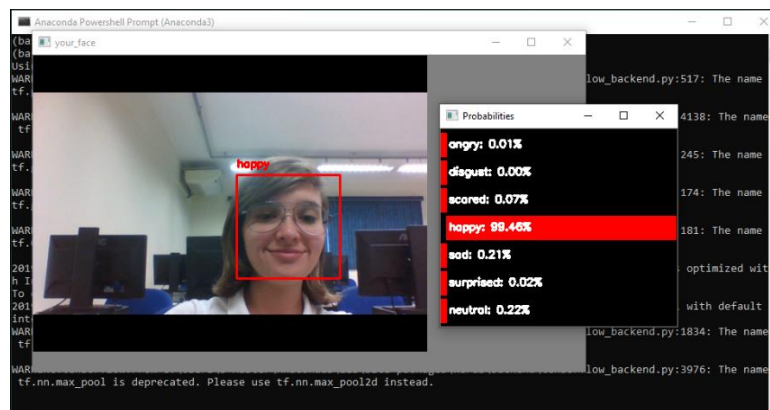


Figura 8: Reconhecimento da face e expressão: Feliz

Fonte: Elaborado pelo autor.

Na Figura 9 é possível visualizar um teste feito em tempo real da rede neural, reconhecendo a expressão de neutra (neutral).

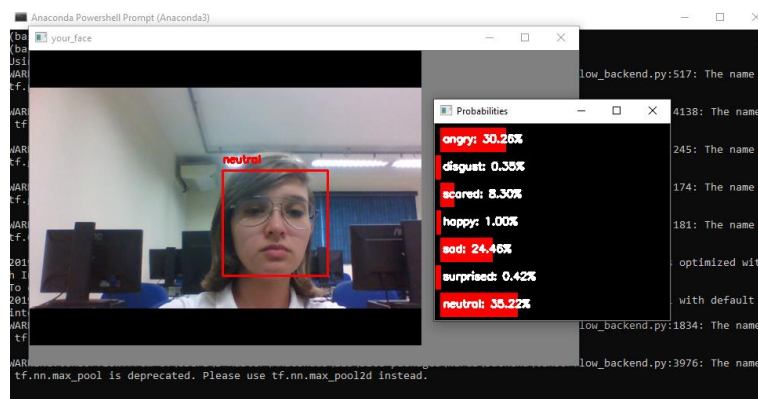


Figura 9: Reconhecimento da face e expressão: Neutra

Fonte: Elaborado pelo autor.

Na Figura 10 é possível visualizar um teste feito em tempo real da rede neural, reconhecendo a expressão de triste (sad).

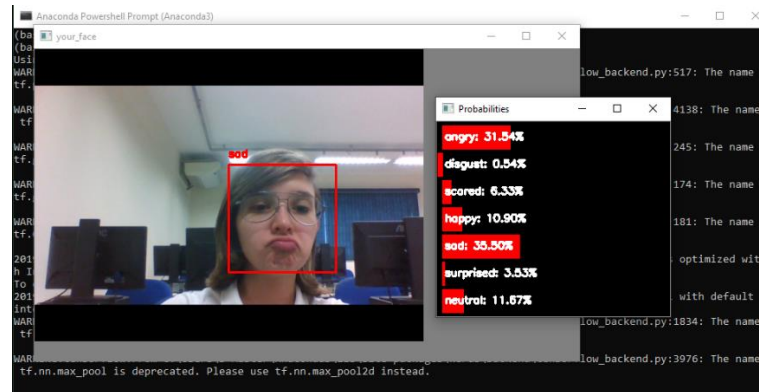


Figura 10: Reconhecimento da face e expressão: Triste

Fonte: Elaborado pelo autor.

Na Figura 11 é possível visualizar um teste feito em tempo real da rede neural, reconhecendo a expressão de assustado (scared).

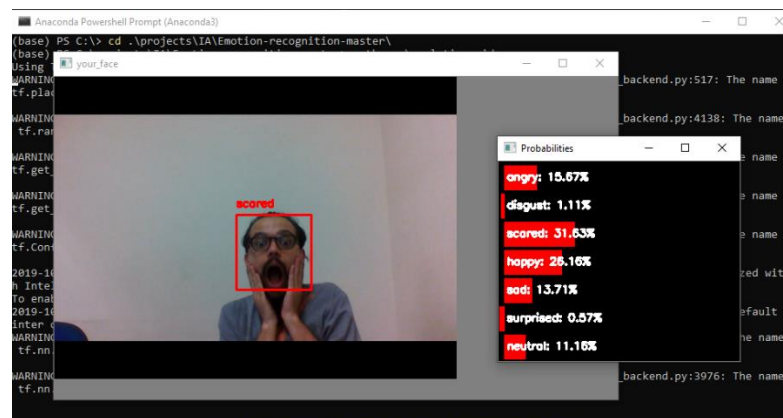


Figura 11: Reconhecimento da face e expressão: Assustado

Fonte: Elaborado pelo autor.

Além da aplicação apresentar as chances de ser uma expressão em tempo real para o usuário visualizar se o que ele está expressando está sendo identificado corretamente, foi criado um arquivo de log, que armazena cada emoção e a probabilidade de ser tal expressão. Esse arquivo serve para realizar um armazenamento dos dados obtidos pelo usuário.

Na Figura 12 é possível visualizar um exemplo do arquivo de log utilizado pela aplicação para armazenar as expressões registradas.

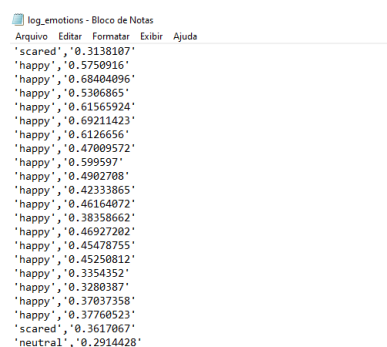


Figura 12: Log do sistema

Fonte: Elaborado pelo autor

Os resultados do reconhecimento das expressões em tempo real por meio da webcam foram analisados e foram fiéis às expressões realizadas pelos usuários usados para teste.

3.4 Análise e apresentação dos resultados

O modelo utilizado tem facilidade para reconhecimento das expressões de felicidade, neutra e raiva. As demais expressões são confusas e a aplicação não a identifica com tanta acurácia.

Para validar a acuracidade da aplicação, foi realizado uma conferência dos *frames* obtidos com a filmagem dos colaboradores em tempo real. A filmagem foi feita a cada hora e a sua duração foi de 2 segundos. Com isso, foi o suficiente para adquirir um total de 2 mil imagens de cada colaborador, no decorrer de uma semana de trabalho (trabalhando 8 horas por dia e 5 dias por semana). Além do armazenamento dos *frames*, a aplicação gerava um arquivo de log com as emoções e a probabilidade de realmente ser aquela emoção. Com ambos, foi realizado um *double check* dos *frames* com os resultados obtidos pelo arquivo de log e assim, gerando gráficos para apresentação dos resultados de acertos e erros da aplicação com cada colaborador, em cada emoção que foi reconhecida.

Deve-se lembrar que existem alguns casos de “falso positivo” e “falso negativo” citados na seção 1.1.2.1.

Na Figura 13 é possível visualizar um exemplo do arquivo de log com a emoção e as chances de ser essa emoção, além do *frame* capturado do colaborador 1.

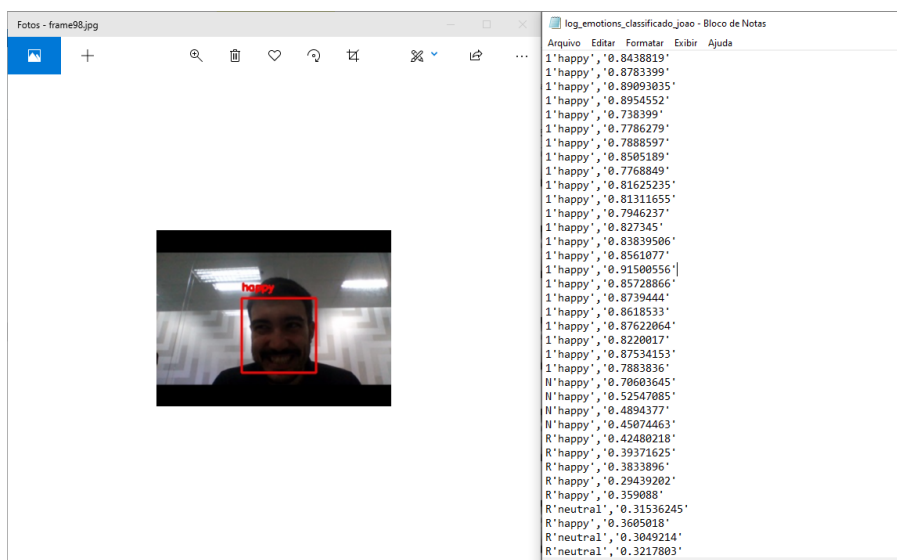


Figura 13: Log do sistema com o *Frame* no ambiente de trabalho, colaborador 1

Fonte: Elaborado pelo autor

Na Figura 14 é possível visualizar um exemplo do arquivo de log com a emoção e as chances de ser essa emoção, além do *frame* capturado do colaborador 2.

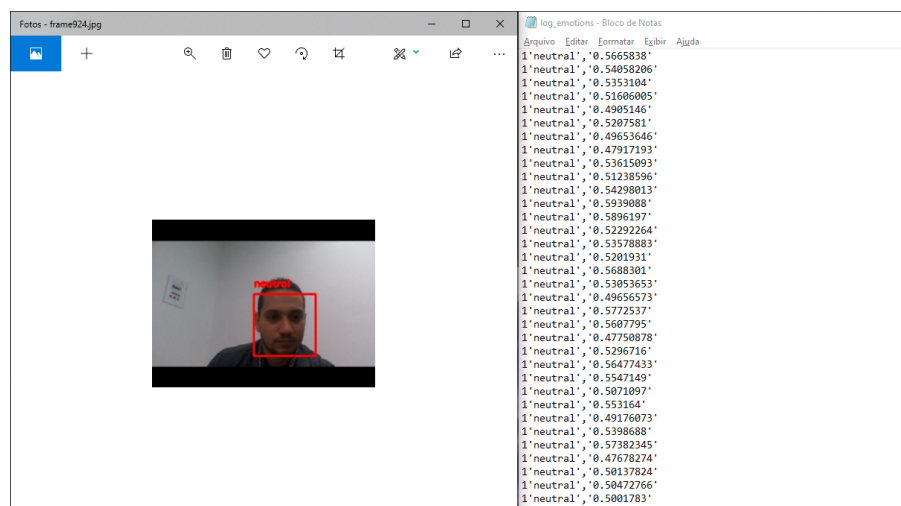


Figura 14: Log do sistema com *Frame* no ambiente de trabalho, colaborador 2

Fonte: Elaborado pelo autor

Na Figura 15 é possível visualizar um exemplo do arquivo de log com a emoção e as chances de ser essa emoção, além do *frame* capturado do colaborador 3.

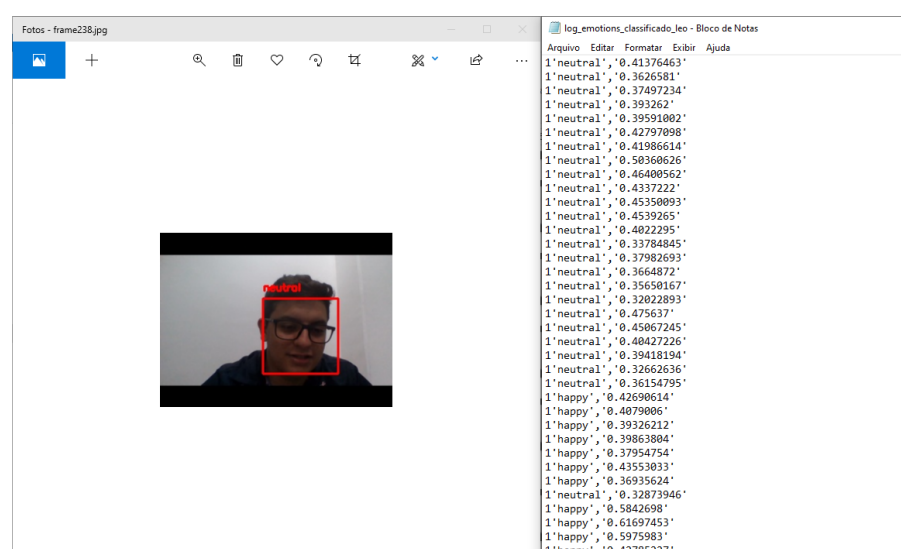


Figura 15: Log do sistema com *Frame* no ambiente de trabalho, colaborador 3

Fonte: Elaborado pelo autor

A partir dos exemplos das figuras acima, é possível validar as expressões faciais realizadas por cada um, com cada *frame* e com o arquivo de log. Com isso, os resultados obtidos pelo arquivo de log foram adicionados à um arquivo do Excel para validação dos acertos e erros. Após essa validação, foi gerado os gráficos a partir dos resultados analisados.

A seguir são apresentados os gráficos com os resultados de cada expressão. Para isso, foi utilizado 2 mil *frames* de cada colaborador e os resultados são com base na quantidade de *frames* que a rede neural convolucional acertou ou errou. A apresentação se inicia com o colaborador 1 e cada expressão capturada de seu rosto, após o colaborador 1 foi realizado o mesmo processo com o colaborador 2 e assim sucessivamente.

Alguns dos valores foram arredondados pelos gráficos.

Verifica-se então que a aplicação ao tentar identificar a expressão de raiva no colaborador 1 acertou em todas as tentativas, acertando 100% das expressões. Na Figura 16 pode-se visualizar esse resultado.

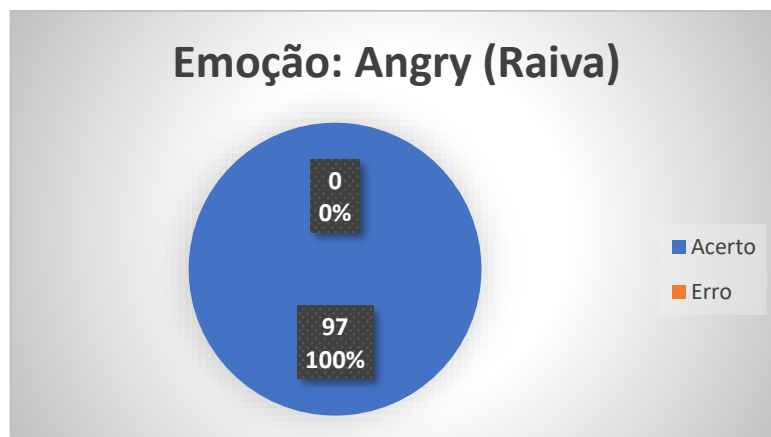


Figura 16: Colaborador 1 - Emoção Raiva

Fonte: Elaborado pelo autor

Validando a próxima expressão que o modelo convolucional identificou como sendo feliz no colaborador 1, o modelo acertou 58,97% das expressões apresentadas e errou 41,02%. Esse resultado é apresentado na Figura 17.

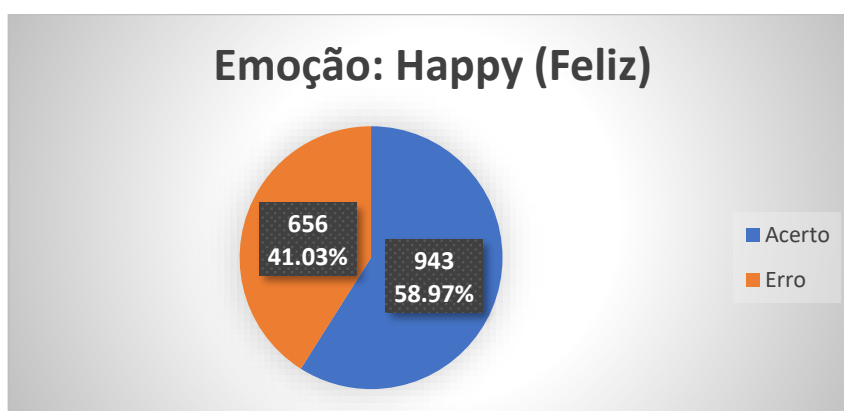


Figura 17: Colaborador 1 - Emoção Feliz

Fonte: Elaborado pelo autor

Verifica-se então que a aplicação ao tentar identificar a expressão neutra no colaborador 1, obteve uma assertividade de 97,26% das expressões apresentadas e com isso, obteve 2,73% de erros com relação as expressões. Com base nessa análise, é possível perceber que o modelo identifica com uma acurácia elevada a expressão neutra. Na Figura 18 é possível avaliar esse resultado.

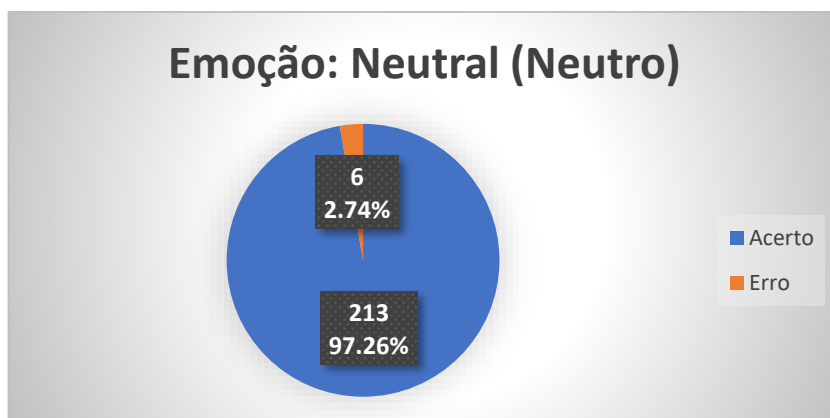


Figura 18: Colaborador 1 - Emoção Neutra

Fonte: Elaborado pelo autor

Analisando os resultados obtidos que o modelo convolucional apresentou ao identificar a expressão triste no rosto do colaborador 1, pode-se verificar que o modelo acertou 18,96% das expressões apresentadas e errou 81,03%. Com isso, percebe-se que o modelo gerou resultados de falso positivo, ou seja, a maioria das expressões realizadas pelo colaborador 1 não eram de tristeza. Na Figura 19 pode-se visualizar esse resultado.

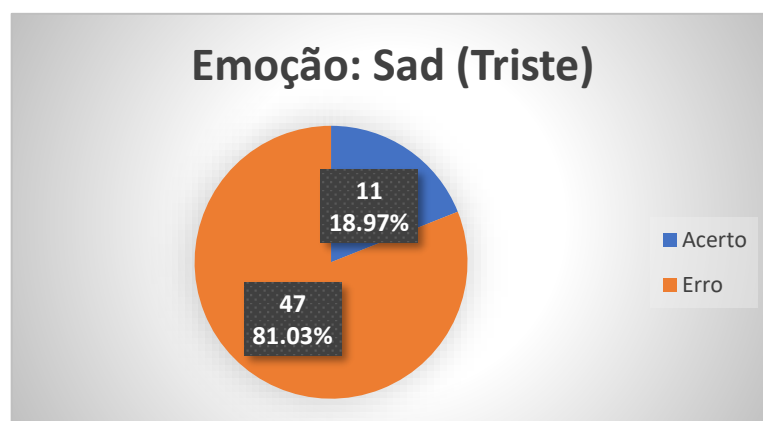


Figura 19: Colaborador 1 - Emoção Triste

Fonte: Elaborado pelo autor

Nota-se que a aplicação ao tentar identificar a expressão de assustado no colaborador 1 acertou 13,04% das expressões apresentadas e errou 86,95%. Analisando esses resultados, o modelo convolucional gerou uma quantidade maior de falsos positivos, pois o a maioria das expressões apresentada pelo colaborador 1 não eram de assustado. Na Figura 20 é apresentado esse resultado.

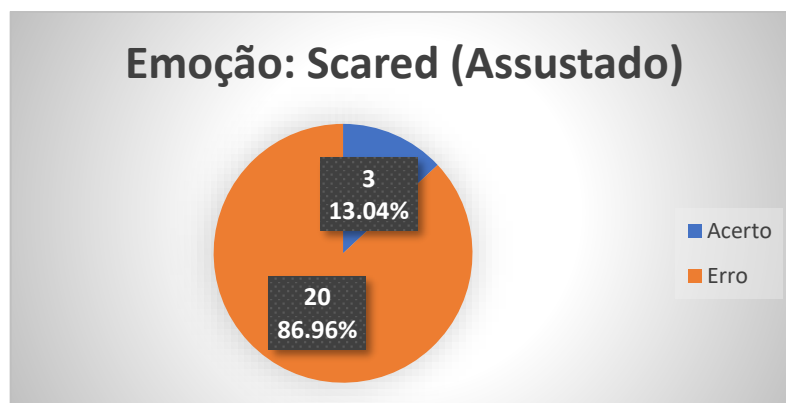


Figura 20: Colaborador 1 - Emoção Assustado

Fonte: Elaborado pelo autor

É possível perceber que a aplicação, ao tentar identificar a expressão de raiva no colaborador 2, acertou 86,27% das expressões apresentadas e errou 13,72%. Na Figura 21 é apresentado esse cenário.

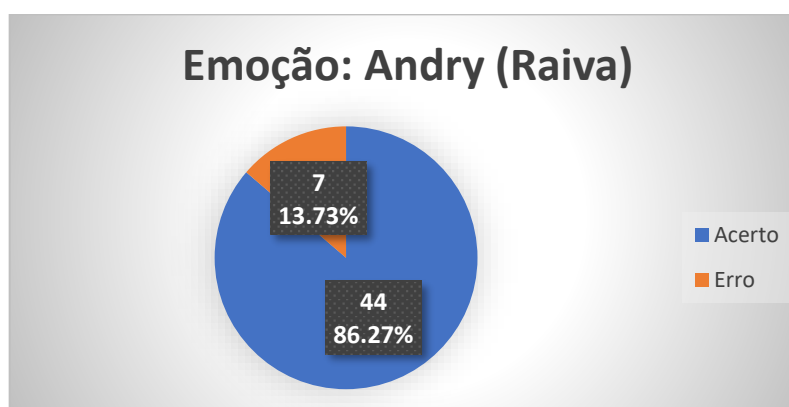


Figura 21: Colaborador 2 - Emoção Raiva

Fonte: Elaborado pelo autor

Com os resultados obtidos pelo modelo convolucional, na identificação da expressão feliz no colaborador 2, obteve-se 60,85% de assertividade nas expressões apresentadas, errando apenas 39,14% das expressões. Na Figura 22 pode-se visualizar esse resultado obtido pelo modelo convolucional.

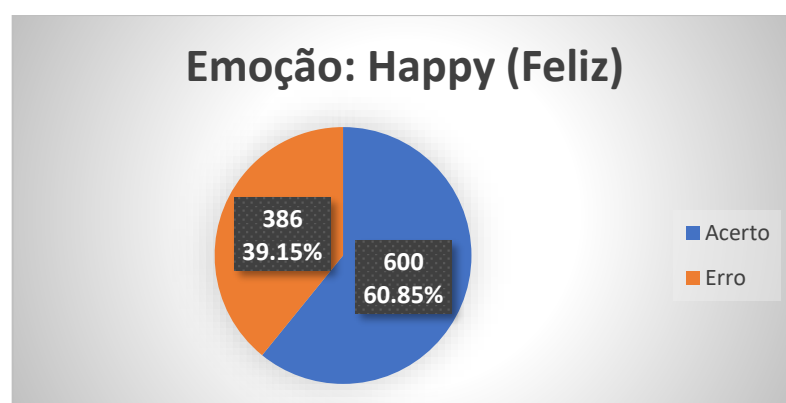


Figura 22: Colaborador 2 - Emoção Feliz

Fonte: Elaborado pelo autor

Validando os resultados obtidos pela aplicação com relação as expressões do colaborador 2, pode-se obter uma assertividade de 99,76% na identificação da expressão neutra e apenas 0,23% de falso positivo. Esses dados são apresentados na Figura 24.

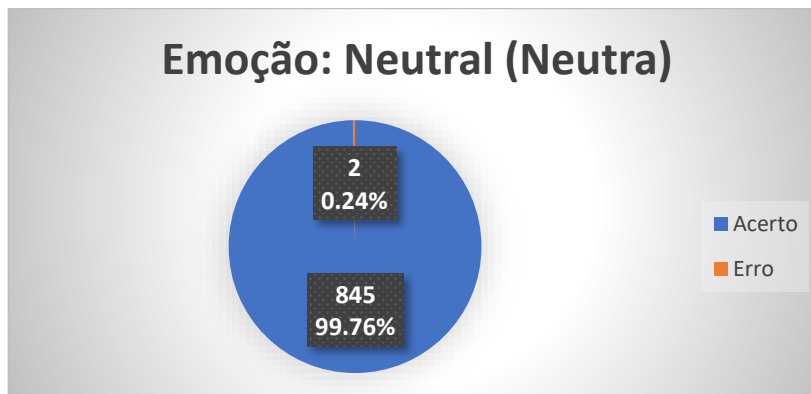


Figura 23: Colaborador 2 - Emoção Neutro

Fonte: Elaborado pelo autor

Avaliando os resultados gerados pelo modelo convolucional na identificação de expressões do colaborador 2, obteve-se 27,5% de acertos e 72,5% de erros. Esses resultados geraram vários falsos positivos, pois o modelo identificava uma expressão que não era realizada pelo colaborador 2. Na Figura 24 apresenta-se esse resultado.

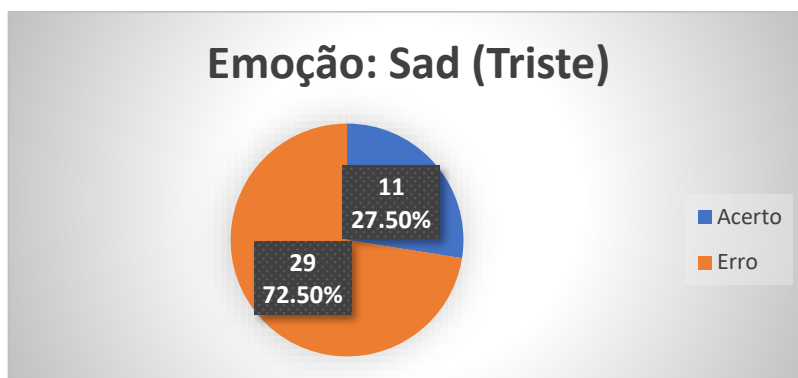


Figura 24: Colaborador 2 - Emoção Triste

Fonte: Elaborado pelo autor

Ao tentar identificar a expressão de assustado no colaborador 2, a aplicação acertou 26,31% das expressões apresentadas e errou 73,68% das expressões realizadas. Na Figura 25 pode-se visualizar esse resultado.

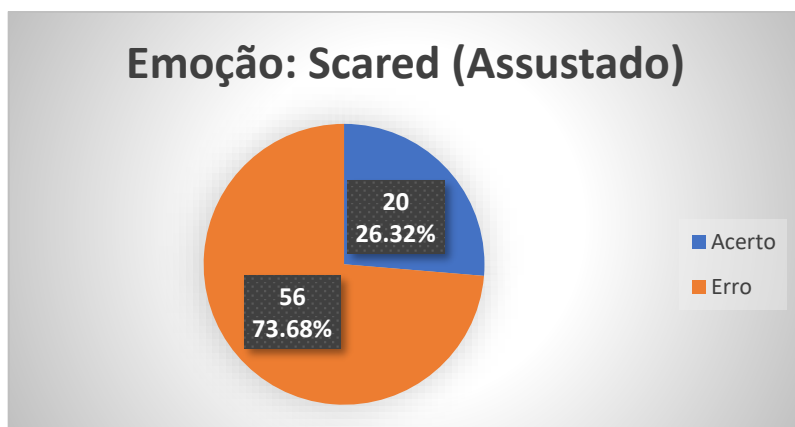


Figura 25: Colaborador 2 - Emoção Assustado
Fonte: Elaborado pelo autor

Verifica-se então que a aplicação ao tentar identificar a expressão de raiva no colaborador 3, acertou 1,92% das expressões apresentadas e errou 98,07% das expressões realizadas (falsos positivos), conforme apresentado na Figura 26.

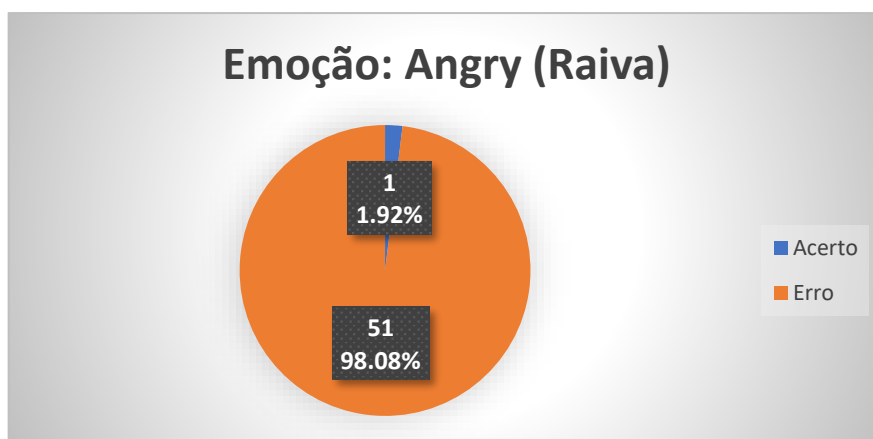


Figura 26: Colaborador 3 - Emoção Raiva
Fonte: Elaborado pelo autor

Analisando os resultados obtidos pela aplicação na identificação da expressão feliz, as expressões realizadas pelo colaborador 3 resultaram em uma assertividade de 60,13% das expressões apresentadas e consequentemente, errou 39,86% das expressões. Na Figura 27 é apresentado esse resultado.

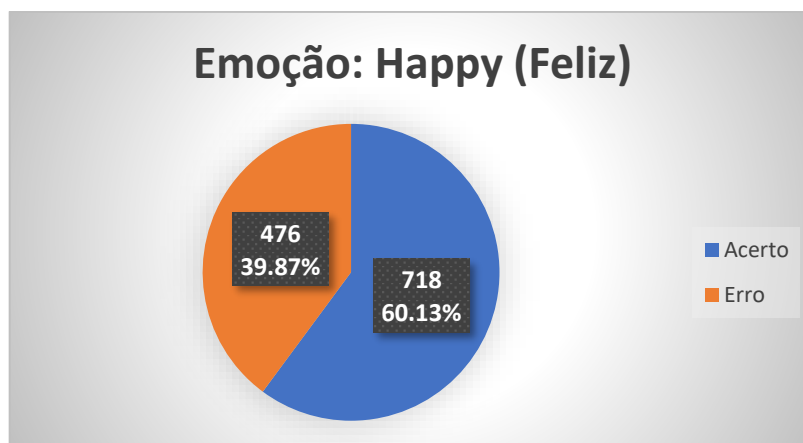


Figura 27: Colaborador 3 - Emoção Feliz

Fonte: Elaborado pelo autor

Percebe-se no levantamento dos resultados que a expressão neutra é identificada com facilidade pelo modelo convolucional. Com isso, ao avaliar as expressões realizadas pelo colaborador 3, o modelo obteve um total de 99,21% de acertos das expressões apresentadas e errou apenas 0,78% . Esses resultados são apresentados na Figura 28.

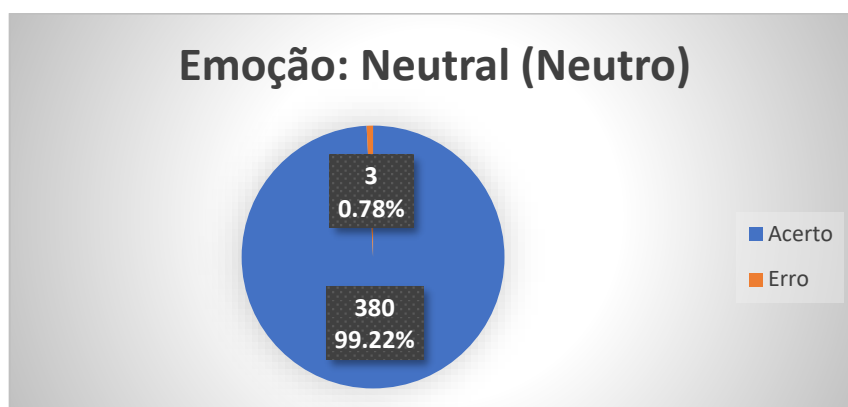


Figura 28: Colaborador 3 - Emoção Neutro

Fonte: Elaborado pelo autor

Nota-se que no processo de identificação da expressão triste, o modelo obteve uma assertividade de 32,43% das expressões e errou 67,56% das expressões realizadas pelo colaborador 3. É possível analisar esse resultado na Figura 29.

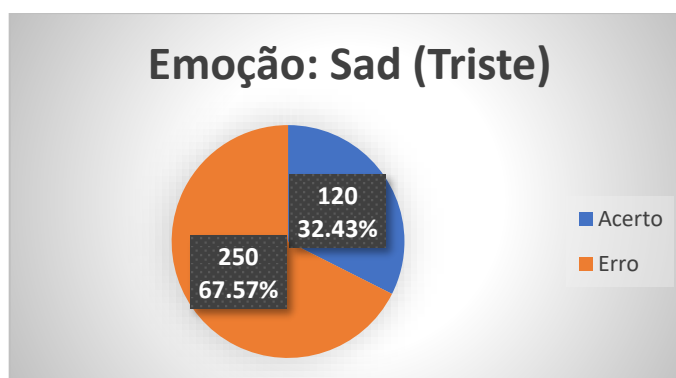


Figura 29: Colaborador 3 - Emoção Triste

Fonte: Elaborado pelo autor

Verifica-se então que a aplicação ao tentar identificar a expressão de assustado no colaborador 3, falhou em todas as tentativas. Analisando as fotos (*frames*) e o arquivo de log, pode-se perceber que o modelo convolucional gerou um falso positivo, ou seja, as expressões realizadas pelo colaborador 3 não eram de assustado. Na Figura 30 pode-se visualizar esse resultado.

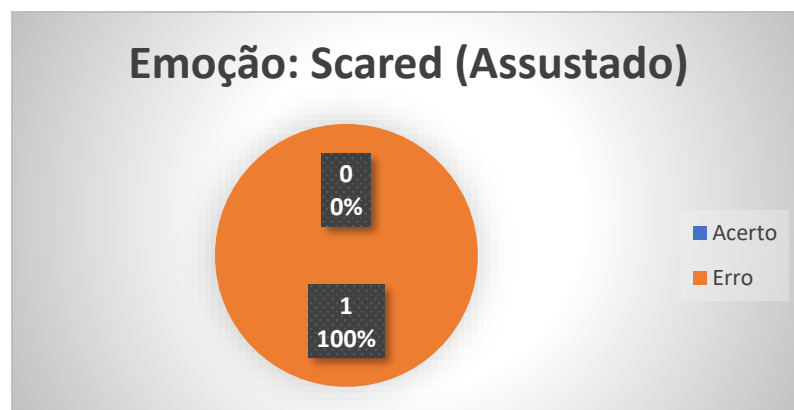


Figura 30: Colaborador 3 - Emoção Assustado

Fonte: Elaborado pelo autor

Após a análise dos resultados e dos gráficos, pode-se perceber que em alguns casos, quando seria acionado a aplicação para captura dos rostos e das expressões, os colaboradores tentaram forçar algumas expressões, dificultando assim a acurácia dos resultados, pois o ideal era validar as expressões naturais de cada colaborador em seu ambiente de trabalho, por isso as expressões de assustado e triste resultaram em muitos falsos positivos.

A acurácia do modelo se torna mais efetiva com as expressões de feliz, neutra e raiva, pois são nessas expressões que o índice de acerto são maiores nos três casos de teste.

Percebeu-se também que, para o colaborador que utilizava um volume maior de barba, a identificação das expressões foi mais difícil, isso pode ter ocorrido devido à uma leve cobertura de sua boca.

Quadro 3: Sintetização dos resultados

Expressão	Acuracidade
Raiva	Em torno de 80% à 90%
Feliz	Em torno de 58% à 61%
Neutro	Em torno de 95% à 99%
Triste	Em torno de 18% à 32%
Surpreso	Em torno de 13% à 18%

Fonte: Elaborado pelo autor.

CONSIDERAÇÕES FINAIS

Este trabalho teve como problema de pesquisa o seguinte questionamento: “Como algoritmos de Inteligência Artificial (IA) podem auxiliar na identificação das expressões faciais em um ambiente empresarial e gerar um relatório sobre o estado aparente de seus colaboradores?” Chega ao seu final com as seguintes considerações:

Para a utilização da aplicação de reconhecimento de expressões faciais e também para tratamento de imagens e vídeos, que no caso foi a rede neural convolucional, que é um tipo de rede neural artificial que utiliza *deep learning* em sua forma de aprendizado, além de uma base de dados contemplando diversas expressões faciais para auxiliar no treinamento do modelo.

Por se tratar da primeira versão da pesquisa, a aplicação foi submetida a uma análise no ambiente empresarial. Durante a utilização da aplicação, foram capturados os *frames* em tempo real das expressões realizadas pelos três colaboradores e armazenados tanto os *frames* quanto um arquivo de log, para assim ser possível a identificação das expressões faciais e analisá-las se realmente o modelo teve uma acurácia satisfatória.

Após a análise dos resultados obtidos nesta pesquisa, percebeu-se que, em alguns casos, quando seria acionado a aplicação para captura dos rostos e das expressões, os colaboradores tentaram forçar algumas expressões, reduzindo assim a acurácia dos resultados, pois o ideal era validar as expressões naturais de cada colaborador em seu ambiente de trabalho, por isso as expressões de assustado e triste resultaram em muitos falsos positivos. Entretanto, a acurácia do modelo se torna mais efetiva com as expressões de feliz, neutra e raiva, pois são nessas expressões que o índice de acerto são maiores nos três casos de teste, atingindo assim o objetivo desta pesquisa, sendo possível a captura das expressões faciais em tempo real pela aplicação e a avaliação das expressões e emoções expressas pelos mesmos.

Essa aplicação de reconhecimento facial, assim como os demais trabalhos científicos, suscita a elaboração de trabalhos futuros, tais como: a possibilidade de melhorar a acuracidade do modelo da rede neural convolucional, para isso talvez seja necessário treinar durante mais épocas ou melhorar o modelo utilizado, além de aprimorar a base de dados utilizada para treinamento da rede neural. Outra ideia, como o modelo realiza a captura dos *frames* em tempo real por meio da filmagem, seria desenvolver seu próprio *dataset* para realizar os testes da aplicação e assim gerar até mesmo um *dataset* dos colaboradores existentes na empresa, melhorando a assertividade nas expressões por serem rostos

conhecidos e vistos com frequência pelo modelo. Um outro trabalho futuro pode ser identificar o meio termo das emoções, por exemplo, um pouco triste, um pouco feliz, muito assustado, muito feliz, etc.

REFERÊNCIAS

ABH – ASSOCIAÇÃO BRASIL HUNTINGTON: **Alteração de Comportamento**. 2016. Disponível em: <<http://abh.org.br/alteracoes-de-comportamento/>> Acesso em: 13 de abril de 2019.

ABAR, C. O CONCEITO "FUZZY". 2004. Disponível em: <<https://www.pucsp.br/~logica/Fuzzy.htm>> Acesso em: 08 de outubro de 2019.

ARRIGA, O., PLÖGER, P. G., VALDENEGRO, M.: **Real-time Convolutional Neural Networks for Emotion and Gender Classification**. 2017. Disponível em: <<https://arxiv.org/pdf/1710.07557.pdf>> Acesso em: 15 de outubro de 2019.

BERGAMINI, C. W. **Psicologia Aplicada a Administração de Empresas: psicologia do comportamento organizacional**. 4ª ed. São Paulo: Atlas, 2010.

BOMBARDELLI, F. G.: Estudo sobre Reconhecimento Facial 2015. Disponível em: <<https://web.inf.ufpr.br/menotti/ci171-2015-2-1/files/final-Felipe-artigo.pdf>> Acesso em: 13 de agosto de 2019.

CHIAVENATO, I. **Recursos Humanos: o capital humano das organizações**. 8. ed. São Paulo: Atlas, 2008.

DAMACENO, S. S., VASCONCELOS, R. O.: **Inteligência Artificial: Uma breve abordagem sobre seu conceito real e o conhecimento popular**. Outubro de 2018. Disponível em: <<https://periodicos.set.edu.br/index.php/cadernoexatas/article/view/5729>> Acesso em: 20 de março de 2019.

DANTAS, A. C., MELO, S., MOURA, F. F., FERNANDES, M. A.: **Reconhecimento dinâmico de emoções através de expressões faciais utilizando árvore de decisão**. Outubro de 2015. Disponível em: <https://www.researchgate.net/publication/283569290_Reconhecimento_dinamico_de_emocoes_atraves_de_expressoes_faciais_utilizando_arvore_de_decisao> Acesso em: 20 de março de 2019.

DANTAS, A. C., GOUVEIRA, E. B., MELO, M. C., GONÇALVES, C. A., NASCIMENTO, M. Z.: **Interface móvel para o reconhecimento de emoções através de expressões faciais em deficientes visuais**. 2018. Disponível em: <<http://www.lipai.facom.ufu.br/admin/papers/2096374473.pdf>> Acesso em: 20 de março de 2019.

Data Science Academy: **Deep Learning Book**. 2019. Disponível em: <<http://www.deeplearningbook.com.br/o-perceptron-parte-1>>. Acesso em 27 de outubro de 2019.

DINIZ, F. A., NETO, F. M. M., JÚNIOR, F. C. L., FONTES, L. M. O.: **RedFace: um sistema de reconhecimento facial baseado em técnicas de análise de componentes principais e auto faces**. 17 de maio de 2013. Disponível em: <<http://seer.upf.br/index.php/rbca/article/view/2627>> Acesso em: 20 de março de 2019.

FERNANDES, J. G. L., SILVA, N. A. M., BROCK, T. R., QUEIROGA, A. P. G., RODRIGUES, L. C.: **Inteligência Artificial: Uma visão geral**. 24 de maio de 2018. Disponível em: <<http://reeced.com.br/index.php/reeced/article/view/25>> Acesso em: 27 de março de 2019.

FURTADO, A. M. M., SOUZA, D. M.: **Descoberta de elenco em trailers de filmes utilizando reconhecimento facial em deep learning**. 2016. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/graduacao/article/view/25672/0>>. Acesso em: 20 de março de 2019.

GÉRON, A.: **Hands-On Machine Learning with Scikit-Learn & TensorFlow - Concepts, Tools, and Techniques To Build Intelligent Systems**. 2017.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GOLDSCHMIDT, R., PASSOS, E.: **Data mining: um guia prático**. Rio de Janeiro: Elsevier, 2005.

GOMIDE, F. A. C.; GUDWIN, R. R.: **Modelagem, Controle, Sistemas E Lógica Fuzzy**. 1994. Universidade Estadual de Campinas (UNICAMP). Disponível em: <<https://pdfs.semanticscholar.org/c4ea/4484bdd2449053d7b7384b05c0e2def86449.pdf>>. Acesso em: 16 março 2019.

GONÇALVES, A.: **Inteligência Artificial e Sistemas Especialista**. 2009. Disponível em: <<https://administradores.com.br/artigos/inteligencia-artificial-e-sistemas-especialistas>>. Acesso em: 03 de novembro de 2019.

GOODFELLOW, I. , BENGIO, Y. e COURVILLE, A.: **Deep Learning**. 2016. Disponível em: <<https://books.google.com.br/books?hl=pt-BR&lr=&id=omivDQAAQBAJ&oi=fnd&pg=PR5&dq=deep+learning+bengio+2016&ots=MT0immBWQ&sig=vwbIo5m2aEZu17WL7i7n5wxXHY0#v=onepage&q=deep%20learning%20bengio%202016&f=false>> Acesso em: 20 de março de 2019.

GUAHYBA, A.: **Visão Geral Sobre Redes Neurais**. 2019. Disponível em: <<http://www.guahyba.vet.br/avicultura/neurais.htm>> Acesso em: 23 de setembro de 2019.

GUIMARÃES, R. M.: **Desenvolvimento de um protótipo de software de reconhecimento facial de tempo real para registro eletrônico de ponto em ambientes indoor com utilização do dispositivo Kinect**. 2015. Disponível em: <<http://www.fumec.br/revistas/sigc/article/view/3037>> Acesso em: 23 de março de 2019.

HAN, J., KAMBER, M.: **Data mining: concepts and techniques**. 2. ed. São Francisco: Morgan Kaufmann, 2006. 770 p.

KERSCHBAUMER, R.: **Sistemas Especialistas**. 2018. Disponível em: <<http://professor.luzerna.ifc.edu.br/ricardo-kerschbaumer/wp-content/uploads/sites/43/2018/02/3-Sistemas-Especialistas.pdf>> Acesso em: 08 de outubro de 2019.

KUMAR, A.: **Achieving 90% accuracy in Object Recognition Task on CIFAR-10 Dataset with Keras: Convolutional Neural Networks**. 2018. Disponível em: <<https://appliedmachinelearning.blog/2018/03/24/achieving-90-accuracy-in-object-recognition-task-on-cifar-10-dataset-with-keras-convolutional-neural-networks/>> Acesso em: 15 de outubro de 2019.

KUMAR, A.: **Demonstration of Facial Emotion Recognition on Real Time Video Using CNN : Python & Keras**. 2018. Disponível em: <<https://appliedmachinelearning.blog/2018/11/28/demonstration-of-facial-emotion-recognition-on-real-time-video-using-cnn-python-keras/>> Acesso em: 09 de setembro de 2019.

LINO, T. A. L. R.: **Inteligência Artificial, Humana e a Emoção**. 2004. Disponível em: <<https://www.psicologia.pt/artigos/textos/TL0012.PDF>>. Acesso 27 de outubro de 2019.

MARSLAND, S.: **Machine Learning: An Algorithmic Perspective** (2ª Edição) 2009.

MONARD, M. C., BARANAUSKAS, J. A.: **Conceitos sobre aprendizado de máquina**. In: REZENDE, S. O. (Org.). **Sistemas inteligentes: fundamentos e aplicações**. Barueri: Manole, 2003. p. 89-114.

PALMIERE, S. E.: **Arquiteturas e Topologias de Redes Neurais Artificiais**. 2016. Disponível em: < <https://www.embarcados.com.br/redes-neurais-artificiais/>> Acesso em: 03 de novembro de 2019.

ROSA, T. O. e LUZ, H. S. **Conceitos Básicos de Algoritmos Genéticos: Teoria e Prática**. 2009. Disponível em: <https://www.researchgate.net/publication/247935438_Conceitos_Basicos_de_Algoritmos_Geneticos_Teoria_e_Pratica> Acesso em: 02 de abril de 2019.

RUSELL, S., NORVIG, P.: **Inteligência Artificial (tradução da segunda edição)** Editora campus. 2003.

SALES, A. I.: **Administração e RH. Módulo I**. 2007. Disponível em: <https://www.cairu.br/biblioteca/arquivos/Administracao/Administracao_RH_Modulo_I.pdf> Acesso 20 de abril de 2019.

SAS Institute Inc.: **Inteligência Artificial. O que é e qual sua importância?**. 2019. Disponível em: <https://www.sas.com/pt_br/insights/analytics/inteligencia-artificial.html> Acesso em: 15 de março de 2019.

SAS Institute Inc.: **Redes Neurais** 2019. Disponível em: <https://www.sas.com/pt_br/insights/analytics/neural-networks.html> Acesso em: 09 de outubro de 2019.

SILVA, J. J.: **Combinação de Classificadores Floresta de Caminhos Ótimos aplicados no Reconhecimento Facial**. Julho de 2016 Disponível em: <<http://www.cc.faccamp.br/Dissertacoes/JairJoseSilva.pdf>> Acesso em: 20 de março de 2019.

SOUSA, A. L., COSTA, S. W. da S., PIRES, Y. P., ARAÚJO, F.: **Reconhecimento de expressões faciais e emocionais como método avaliativo de aplicações computacionais**. 2016. Disponível em: <<http://deposita.ibict.br/handle/deposita/18>> Acesso em: 20 de março de 2019.

STREHL, M.: **Sistemas Especialistas – Definições e Exemplos**. 2000. Disponível em: <<http://www.inf.ufsc.br/~j.barreto/trabaluno/IASteel001.pdf>> Acesso em: 02 de abril de 2019.

TAIGMAN, Y., YANG, M., RANZATO, M. A., WOLF L.: **DeepFace: Closing the Gap to Human-Level Performance in Face Verification**. 2014. Disponível em: <https://www.cs.toronto.edu/~ranzato/publications/taigman_cvpr14.pdf> Acesso em 27 de agosto de 2019.

TEIXEIRA, A. R. S.: **Reconhecimento de padrões nas expressões faciais dos usuários através do uso da mineração de dados**. 13 de abril de 2018. Disponível em: <<http://bdm.ufpa.br/jspui/handle/prefix/1044>> Acesso em: 16 de março de 2019.

VASCONCELLOS, P.: **Como saber se o seu modelo de machine learning está funcionando mesmo**. 2018. Disponível em: <<https://paulovasconcellos.com.br/como-saber-se-seu-modelo-de-machine-learning-est%C3%A1-funcionando-mesmo-a5892f6468b>> Acesso em: 30 de outubro de 2019.

WANGENHEIM, C. G. V.: **Raciocínio Baseado em Casos**. 2003. Disponível em: <https://books.google.com.br/books?hl=pt-BR&lr=&id=cIMwqPOdPTMC&oi=fnd&pg=PA1&dq=racioc%C3%ADnio+baseado+em+casos&ots=RuWRU9Obti&sig=2i-Na_1P2_TneHSnOqZPFATZAaw#v=onepage&q=racioc%C3%ADnio%20baseado%20em%20casos&f=false> Acesso em: 08 de outubro de 2019.

WOYCIEKOSKI, C., HUTZ, C. S.: **Inteligência emocional: teoria, pesquisa, medida, aplicações e controvérsias**. 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-79722009000100002> Acesso em: 09 de agosto de 2019.

ZANELATO, A. L., FERREIRA, C. B.: **Sistema de reconhecimento de expressões faciais**. 2017. Disponível em: <http://repositorio.roca.utfpr.edu.br/jspui/bitstream/1/8459/1/CT_COENC_2017_1_6.pdf> Acesso em: 20 de março de 2019.

ANEXOS

MODELO DE TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

Neste ato, Lucas Fagundes de Campos, nacionalidade Brasileira, estado civil Solteiro, portador da Cédula de identidade RG nº 40.147.783-6. AUTORIZO o uso de minha imagem em todo e qualquer material entre fotos e documentos, para ser utilizada no Trabalho de Graduação – Inteligência Artificial para Reconhecimento de Emoções em um Ambiente Empresarial, do curso de Análise e Desenvolvimento de Sistemas, com sede na FATEC de Indaiatuba - Dr Archimedes Lammoglia, sejam essas destinadas à divulgação ao público em geral. A presente autorização é concedida a título gratuito, abrangendo o uso da imagem acima mencionada em todo território nacional e no exterior, das seguintes formas: (I) trabalho escrito; (II) busdoor; folhetos em geral (encartes, mala direta, catálogo, etc.); (III) folder de apresentação; (IV) mídia eletrônica (painéis, vídeo-tapes, televisão, cinema, programa para rádio, entre outros). Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser reclamado a título de direitos conexos à minha imagem ou a qualquer outro, e assino a presente autorização em 02 vias de igual teor e forma.

Indaiatuba, dia 11 de Dezembro de 2019.

Lucas F. Campos

(assinatura)

Nome:

Telefone p/ contato: 19 9 9988-1993

MODELO DE TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

Neste ato, José Roberto de Valle Simões, nacionalidade brasileira, estado civil solteiro, portador da Cédula de identidade RG nº. 46241081-4. AUTORIZO o uso de minha imagem em todo e qualquer material entre fotos e documentos, para ser utilizada no Trabalho de Graduação – Inteligência Artificial para Reconhecimento de Emoções em um Ambiente Empresarial, do curso de Análise e Desenvolvimento de Sistemas, com sede na FATEC de Indaiatuba - Dr Archimedes Lammoglia, sejam essas destinadas à divulgação ao público em geral. A presente autorização é concedida a título gratuito, abrangendo o uso da imagem acima mencionada em todo território nacional e no exterior, das seguintes formas: (I) trabalho escrito; (II) busdoor; folhetos em geral (encartes, mala direta, catálogo, etc.); (III) folder de apresentação; (IV) mídia eletrônica (painéis, vídeo-tapes, televisão, cinema, programa para rádio, entre outros). Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser reclamado a título de direitos conexos à minha imagem ou a qualquer outro, e assino a presente autorização em 02 vias de igual teor e forma.

Indaiatuba, dia 11 de Dezembro de 2019.

José Roberto de Valle Simões

(assinatura)

Nome: José Roberto de Valle Simões

Telefone p/ contato: (12) 99793-0208

MODELO DE TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

Neste ato, Leonardo Garcia Bombassi, nacionalidade brasileiro, estado civil solteiro, portador da Cédula de identidade RG nº. 45.887.809-1. AUTORIZO o uso de minha imagem em todo e qualquer material entre fotos e documentos, para ser utilizada no Trabalho de Graduação – Inteligência Artificial para Reconhecimento de Emoções em um Ambiente Empresarial, do curso de Análise e Desenvolvimento de Sistemas, com sede na FATEC de Indaiatuba - Dr Archimedes Lammoglia, sejam essas destinadas à divulgação ao público em geral. A presente autorização é concedida a título gratuito, abrangendo o uso da imagem acima mencionada em todo território nacional e no exterior, das seguintes formas: (I) trabalho escrito; (II) busdoor; folhetos em geral (encartes, mala direta, catálogo, etc.); (III) folder de apresentação; (IV) mídia eletrônica (painéis, vídeo-tapes, televisão, cinema, programa para rádio, entre outros). Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser reclamado a título de direitos conexos à minha imagem ou a qualquer outro, e assino a presente autorização em 02 vias de igual teor e forma.

Indaiatuba, dia 11 de dezembro de 2019.

Leonardo Bombassi

(assinatura)

Nome: Leonardo Garcia Bombassi
Telefone p/ contato: (19) 99943-7993

MODELO DE TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

Neste ato, Mateus Rocha Pereira, nacionalidade Brasileira, estado civil casado, portador da Cédula de identidade RG nº. 32.507742-0. AUTORIZO o uso de minha imagem em todo e qualquer material entre fotos e documentos, para ser utilizada no Trabalho de Graduação – Inteligência Artificial para Reconhecimento de Emoções em um Ambiente Empresarial, do curso de Análise e Desenvolvimento de Sistemas, com sede na FATEC de Indaiatuba - Dr Archimedes Lammoglia, sejam essas destinadas à divulgação ao público em geral. A presente autorização é concedida a título gratuito, abrangendo o uso da imagem acima mencionada em todo território nacional e no exterior, das seguintes formas: (I) trabalho escrito; (II) busdoor; folhetos em geral (encartes, mala direta, catálogo, etc.); (III) folder de apresentação; (IV) mídia eletrônica (painéis, vídeo-tapes, televisão, cinema, programa para rádio, entre outros). Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser reclamado a título de direitos conexos à minha imagem ou a qualquer outro, e assino a presente autorização em 02 vias de igual teor e forma.

Indaiatuba, dia 11 de dezembro de 2019.

Mateus R. Pereira
(assinatura)

Nome: Mateus Rocha Pereira
Telefone p/ contato: (19) 98149-1924