TECNOLOGIA BIOMÉTRICA DE RECONHECIMENTO PESSOAL

TECNOLOGÍA BIOMÉTRICA DE RECONOCIMIENTO PERSONAL

BIOMETRIC TECHNOLOGY OF PERSONAL RECOGNITION

Alla KAPITON[1]
e-mail: kits_seminar@ukr.net
Nataliia KONONETS[2]
e-mail: natalkapoltava7476@gmail.com
Volodymyr MOKLIAK[3]
e-mail: vovchik01071981@gmail.com
Valentyna ONIPKO[4]
e-mail: valentyna.onipko@pdau.edu.ua
Serhiy DUDKO[5]
e-mail: dudko@pano.pl.ua
Vadym PYLYPENKO[6]
e-mail: pylypenko@pano.pl.ua
Anna SOKIL[7]
e-mail: sokol7227@ukr.net

Como referenciar este artigo:

KAPITON, A.; KONONETS, N.; MOKLIAK, V.; ONIPKO, V.; DUDKO, S.; PYLYPENKO, V.; SOKIL, A. Tecnologia biométrica de reconhecimento pessoal. **Revista on line de Política e Gestão Educacional**, Araraquara, v. 28, n. 00, e023015, 2024. e-ISSN: 1519-9029. DOI: https://doi.org/10.22633/rpge.v28i00.19390

| **Submetido em**: 11/03/2024
| **Revisões requeridas em:** 08/04/2024
| **Aprovado em**: 16/05/2024
| **Publicado em**: 19/06/2024

Editor: Prof. Dr. Sebastião de Souza Lemes
Editor Adjunto Executivo: Prof. Dr. José Anderson Santos Cruz

[1] Universidade Nacional "Yuri Kondratyuk Poltava Politécnica", Poltava – Ucrânia. Professor Associado do Departamento de Tecnologias e Sistemas de Informação e Computação. Doutor em Ciências Pedagógicas.
[2] Universidade Cooperativa "Poltava de Economia e Comércio", Poltava – Ucrânia. Professor Associado do Departamento de Cibernética Econômica, Economia de Negócios e Sistemas de Informação. Doutor em Ciências Pedagógicas.
[3] Universidade Pedagógica Nacional V. G. Korolenko de Poltava, Poltava – Ucrânia. Professor do Departamento de Pedagogia Geral e Andragogia. Doutor em Ciências Pedagógicas.
[4] Universidade Estadual Agrária de Poltava, Poltava – Ucrânia. Professor do Departamento de Agricultura e Agroquímica nomeado após V. I. Sazanov, Departamento de Construção e Educação Profissional. Doutor em Ciências Pedagógicas.
[5] Academia de Educação Contínua M. V. Ostrogradsky de Poltava, Poltava – Ucrânia. Diretor Adjunto. Doutor em Ciências Pedagógicas.
[6] Academia de Educação Contínua M. V. Ostrogradsky de Poltava, Poltava – Ucrânia. Primeiro Diretor Adjunto. Doutor em Ciências Pedagógicas.
[7] Universidade Pedagógica Nacional V. G. Korolenko de Poltava, Poltava – Ucrânia. Estudante de Pós-Graduação do Departamento de Pedagogia Geral e Andragogia.

RESUMO: Atualmente, o processamento de imagens é amplamente utilizado em sistemas de segurança para reconhecer pessoas. Para este estudo, foram selecionados algoritmos de aprendizado de máquina, considerando a disponibilidade limitada de dados. O trabalho analisou a área de reconhecimento facial, a relevância deste sistema nos dias de hoje, o reconhecimento biométrico do sistema «*Face ID*», diversos métodos para o reconhecimento facial e investigou em quais áreas de atividade o sistema de reconhecimento facial é utilizado e para qual finalidade. Como parte do estudo, diversos algoritmos de reconhecimento facial foram analisados. Foi comprovado que todo o sistema de reconhecimento facial pode ser modelado utilizando o método de extração de contornos Violy-Jones e testado com um resultado bem-sucedido de reconhecimento de aproximadamente 75%. As capacidades funcionais da biblioteca de visão computacional *OpenCV* e outras bibliotecas foram consideradas.

PALAVRAS-CHAVE: Sistema de reconhecimento de objetos. Algoritmo. Visão computacional. Tecnologias biométricas.

RESUMEN: *En la actualidad, el procesamiento de imágenes se utiliza ampliamente en los sistemas de seguridad para reconocer personas. Para este estudio se seleccionaron algoritmos de aprendizaje automático, en presencia de una cantidad limitada de datos. El trabajo analizó el área temática del reconocimiento facial, la relevancia de este sistema en nuestro tiempo, el reconocimiento biométrico del sistema "FaceID", varios métodos diferentes para el reconocimiento facial, e investigó en qué áreas de actividad se utiliza el sistema "Face recognition" y con qué propósito. Como parte del estudio, se analizaron varios algoritmos de reconocimiento facial. Se ha comprobado que todo el sistema de reconocimiento de caras puede modelarse utilizando el método de extracción de contornos Violy-Jones y se ha probado con un resultado de reconocimiento satisfactorio de aproximadamente el 75%. Se consideran las capacidades funcionales de la biblioteca de visión por ordenador OpenCV y de otras bibliotecas.*

PALABRAS CLAVE: *Sistema de reconocimiento de objetos. Algoritmo. Visión por computadora. Tecnologías biométricas.*

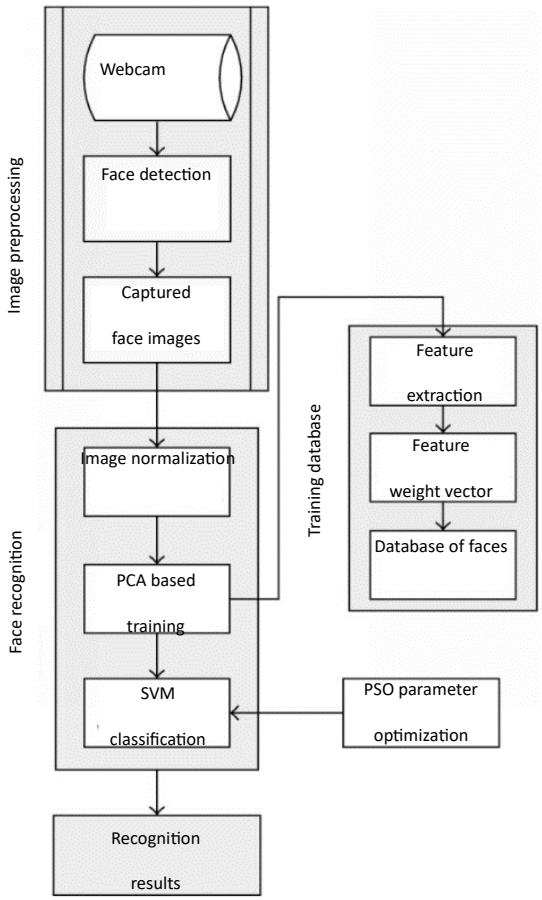
ABSTRACT: *Today, image processing is widely used in security systems to recognize people. For this study, machine learning algorithms were selected, in the presence of a limited amount of data. The work analyzed the subject area of face recognition, the relevance of this system in our time, biometric recognition of the «FaceID» system, several different methods for face recognition, and investigated in which areas of activity the «Face recognition» system is used for which purpose. As part of the study, a number of face recognition algorithms were analyzed. It has been proven that the entire face recognition system can be modeled using the Violy-Jones contour extraction method and tested with a successful recognition result of approximately 75%. The functional capabilities of the OpenCV computer vision library and other libraries are considered.*

KEYWORDS: *Object recognition system. Algorithm. Computer vision. Biometric technologies.*

Introdução

Os principais conceitos do sistema de reconhecimento facial envolvem a capacidade de detectar e identificar uma ou mais pessoas em uma imagem, ou sequência de quadros de vídeo. Um sistema tecnológico dedicado a essa função é denominado sistema de reconhecimento facial. Esses sistemas utilizam diversos métodos de identificação para comparar uma imagem fornecida com rostos presentes em uma base de dados. Apesar de sua precisão relativamente menor em comparação com tecnologias como reconhecimento de íris e impressões digitais, o reconhecimento facial é amplamente empregado em sistemas de segurança devido à sua capacidade de identificação sem contato direto (Engelbrecht, 2007). O processo geral de reconhecimento facial pode ser visualizado no esquema apresentado na Figura 1.

Figura 1 – Esquema do processo de reconhecimento facial



Fonte: Elaborado pelos autores.

O reconhecimento facial compreende duas etapas principais. A primeira etapa consiste na seleção de características faciais, enquanto a segunda etapa classifica os objetos

identificados. Cada rosto humano possui numerosos pontos nodais que são utilizados como características distintivas, tais como a distância entre os olhos, a largura do nariz, a profundidade das cavidades oculares, o formato das maçãs do rosto e o comprimento da linha da mandíbula. Esses pontos nodais são medidos para gerar um código numérico denominado "*faceprint*", que representa o rosto na base de dados (Engelbrecht, 2007).

No passado, o *software* de reconhecimento facial dependia de imagens 2D para comparar ou identificar outras imagens 2D em uma base de dados. Para ser eficaz e preciso, a imagem precisava capturar o rosto quase diretamente voltado para a câmera, com pouca variação na luz ou na expressão facial em relação à imagem na base de dados. Isso representava um desafio significativo, pois muitas vezes as fotografias não eram tiradas em ambientes controlados. Mesmo pequenas variações na iluminação ou na orientação podiam degradar o desempenho do sistema, resultando em dificuldades para encontrar correspondências precisas na base de dados e levando a altas taxas de erro (Engelbrecht, 2007).

Anteriormente, os principais usuários de *software* de reconhecimento facial eram agências de aplicação da lei, que utilizavam esses sistemas para identificar pessoas em multidões. Algumas agências governamentais também empregam esses sistemas para garantir a segurança e prevenir fraudes eleitorais. Um exemplo é o programa US-VISIT (*US Visitor and Immigrant Status Determination Technology*) do governo dos EUA, que aplica essa tecnologia aos visitantes estrangeiros admitidos no país. Quando um turista estrangeiro recebe um visto, suas impressões digitais e uma fotografia são registradas. Esses dados são então verificados em um banco de dados de criminosos conhecidos e suspeitos de terrorismo. Quando os viajantes chegam a um ponto de verificação nos EUA, suas impressões digitais e fotografias são novamente utilizadas para garantir sua identidade (Engelbrecht, 2007).

Pesquisas científicas exploram os problemas relacionados ao reconhecimento facial com base em momentos invariantes (Reinartz, 2002; Jankowski; Grochowski, 2004). Os momentos invariantes são frequentemente utilizados como características em tarefas de identificação de pessoas. O trabalho de Jankowski e Grochowski (2004) investiga as propriedades desses momentos, demonstrando que diferentes dados apresentam sensibilidades variadas a mudanças. Apesar das abordagens existentes para seleção de características faciais, como lábios, nariz e perfil facial, considerando diversos fatores complicadores na análise de imagens (como ruído, variações na orientação facial e expressões emocionais), ainda não existe uma abordagem universalmente eficaz para resolver esses desafios.

Propõe-se uma abordagem combinada que integra diversas ferramentas para o

reconhecimento facial. Isso inclui métodos de seleção de momentos invariantes, formação de classes de referência de pessoas, a métrica Euclidiana-Mahalanobis, e o uso de redes neurais artificiais (Reinartz, 2002). Transformações generalizadas modificadas para processamento de imagens tridimensionais com parâmetros de rotação e escala desconhecidos têm sido abordadas em vários estudos (Hart, 1968). Os resultados dos algoritmos para detecção rápida de objetos utilizando o método de cascata de funções simples e métodos para reconhecimento facial em tempo real são discutidos (Gates, 1972; Aha; Kibler; Albert, 1991).

Entre os estudos recentes dedicados ao reconhecimento facial, destacam-se os trabalhos de Wilson (1972) e Kibler e Aha (1987), que desenvolvem modelos teórico-probabilísticos para imagens em meio-tom e aplicam métodos de identificação de pessoas com base na regra Bayesiana. Tomek (1976) enfatiza a importância de pesquisar sistemas modernos de reconhecimento facial com base em estudos aprofundados de problemas de aprendizado de máquina.

O estudo das questões relacionadas às ferramentas de reconhecimento de imagem por computador e sua visualização com base em dados recebidos inclui a construção de histogramas de gradientes orientados para identificação de pessoas (Jankowski, 2000). A segmentação eficaz de imagens com base em sua representação gráfica é explorada por Broadley (1993) e Wilson (2000). O uso de redes neurais para reconhecimento facial é amplamente investigado no trabalho de *et al*. (1975).

Os principais métodos e técnicas de representação e reconhecimento de gestos, bem como os métodos para reconhecimento facial dinâmico utilizando estudos de conteúdo de vídeo, estão definidos (Skalak, 1994; Domingo; Gavalda; Watanabe, 1999). Foram estabelecidas as principais etapas de projeto para o desenvolvimento e implementação de sistemas de vigilância e monitoramento por vídeo (Kohonen, 1988; Madigan *et al*., 2002). Problemas relacionados ao reconhecimento de atividades visuais e interações usando análise estocástica, incluindo a projeção de modelos de imagem de fundo adaptativos para detecção e rastreamento de pessoas em tempo real, são abordados em diversos estudos científicos (Suykens; Vandewalle, 1999; Reeves; Bush, 2001; Li *et al.,* 2007; Sane; Ghatol, 2007; Evans, 2008; Koskimaki *et al.*, 2008; Subbotin, 2013a, 2013b).

Atualmente, há uma crescente popularização de *softwares* em diversas aplicações, à medida que os sistemas se tornam mais acessíveis e seu uso se generaliza. Esses sistemas já estão integrados com câmeras e computadores utilizados em bancos, aeroportos e outros ambientes sociais. A *Transport Security Administration* (TSA) dos Estados Unidos está

atualmente testando um software para passageiros frequentes, que precisam se registrar na plataforma. Esse software permite uma triagem rápida dos passageiros e realiza uma avaliação de ameaças à segurança. Para otimizar o fluxo nos aeroportos, as filas foram organizadas em duas colunas, onde uma delas utiliza biometria facial para verificação dos passageiros.

Outros aplicativos integram funcionalidades da TSA e serviços de transferência de dinheiro. Esses programas conseguem verificar rapidamente a identidade do cliente, mediante a autorização para armazenar uma imagem digital. O *software FaceID* gera uma impressão facial a partir dessa imagem para proteger os clientes contra roubo de identidade. O uso de reconhecimento facial elimina a necessidade de apresentar documento de identidade com foto, cartão bancário ou Número Pessoal de Identificação (PIN) para verificar a identidade do cliente, auxiliando as empresas na prevenção de fraudes.

Pelo menos todos os exemplos mencionados acima funcionam com permissão humana; no entanto, existem também sistemas que não exigem autorização. Às vezes, o sistema pode fotografar o cliente sem permissão, violando assim a lei de privacidade. É evidente que essas tecnologias frequentemente têm sido criticadas, pois não se sabe como podem afetar a vida de uma pessoa. Os danos causados pelas tecnologias de reconhecimento facial à privacidade, à liberdade de expressão e ao devido processo afetam a todos nós e não devem ser tratados levianamente. Mesmo que a tecnologia de reconhecimento facial não apresentasse problemas de precisão enviesada ou fosse implantada de forma aleatória e descuidada, aumentando a probabilidade de erro, ainda representaria uma séria ameaça aos valores democráticos, funcionando exatamente como pretendido.

Este estudo investiga os processos de reconhecimento e processamento de imagens amplamente utilizados em modernos sistemas de informação de segurança para o reconhecimento de pessoas. Durante a execução das tarefas atribuídas, foi realizada uma análise detalhada do reconhecimento facial biométrico usando o sistema *FaceID*, explorando diferentes métodos de reconhecimento facial e investigando seu escopo de implementação e áreas de aplicação. Uma contribuição significativa da pesquisa foi o desenvolvimento e avaliação de um sistema de reconhecimento facial baseado no método de seleção de contorno Viola-Jones, considerado eficaz e adequado para implementação. Os autores dedicaram especial atenção à funcionalidade da biblioteca de visão computacional *OpenCV*, destacando a necessidade de futuras investigações aprofundadas.

Materiais e métodos

O objetivo deste estudo é determinar a viabilidade de implementar um sistema de reconhecimento facial utilizando visão computacional e bibliotecas de reconhecimento facial. As tarefas de pesquisa incluem o estudo da tecnologia de reconhecimento facial, análise de métodos de detecção, reconhecimento, formação de pontos de características e codificação de características faciais com base em um retrato eletrônico. O resultado planejado deste trabalho é o desenvolvimento de um sistema capaz de identificar uma pessoa através de seu rosto humano.

Entre os métodos de reconhecimento biométrico mais promissores, destaca-se o reconhecimento facial. Essa técnica oferece diversas vantagens sobre métodos similares, incluindo alta precisão na identificação, capacidade de verificação remota, análise anônima e exigência apenas de uma câmera de vídeo. A diversidade de algoritmos disponíveis, aliada à velocidade e precisão na busca, possibilita que o sistema opere eficazmente em diversas condições. Essas características impulsionaram o desenvolvimento do método, tornando-o o segundo mais comum após a impressão digital.

Para melhorar ainda mais a precisão, uma combinação de múltiplos algoritmos de análise facial é essencial. Por exemplo, a identificação da orelha complementa a identificação facial com eficácia comprovada. Contudo, a otimização inadequada ao usar múltiplos algoritmos pode neutralizar os benefícios dessa abordagem combinada. Uma das tendências mais promissoras no mercado de biometria é a introdução de câmeras digitais inteligentes com capacidade de análise facial integrada. Essas câmeras não apenas oferecem alta qualidade de imagem, mas também podem anexar metadados às imagens, contendo informações sobre os rostos detectados. Isso reduz a carga de *hardware*, diminuindo o custo dos sistemas de reconhecimento biométrico e tornando-os mais acessíveis. Além disso, a transmissão de dados comprimidos e um pequeno fluxo de imagens de detecção de rostos aliviam os canais de dados.

A escolha do método Viola-Jones pelos pesquisadores deste estudo baseia-se em sua eficiência comprovada na busca de objetos em imagens e vídeos em tempo real. Estudos realizados por cientistas nacionais e internacionais demonstram que esse método apresenta baixa probabilidade de erro na identificação de pessoas. A precisão de reconhecimento utilizando este método é consideravelmente alta, o que é um resultado desejável. No entanto, é importante notar que o método padrão não é capaz de detectar rostos humanos virados em ângulos arbitrários, o que pode limitar sua aplicação em sistemas de produção modernos, considerando as crescentes exigências tecnológicas (Winarno *et al.,* 2018).

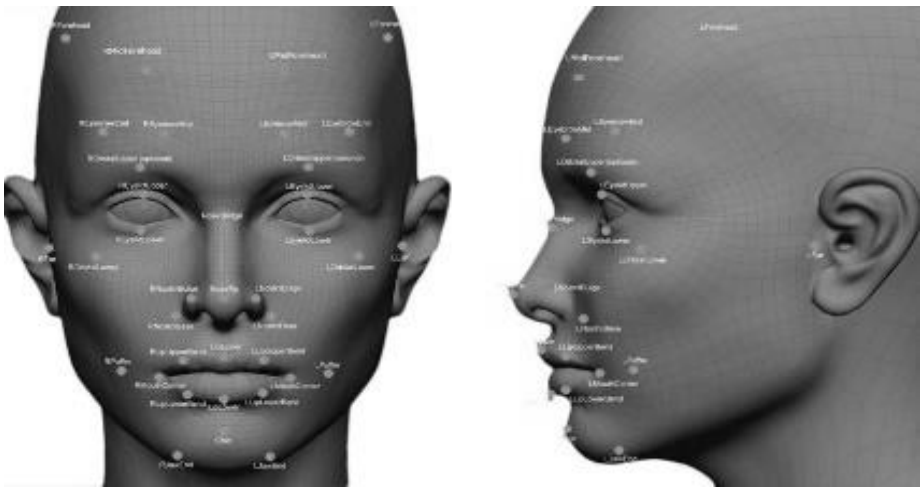
Sistemas de reconhecimento facial são utilizados amplamente não apenas para fins como identificação de criminosos ou pessoas procuradas em locais movimentados, mas também para tarefas domésticas cotidianas. Com a proliferação de câmeras e a melhoria contínua dos algoritmos de detecção e análise facial, a precisão do reconhecimento aumentou significativamente. Apesar das semelhanças funcionais entre os softwares disponíveis, a escolha dos usuários geralmente é baseada em suas preferências individuais.

Resultados e Discussão

O reconhecimento facial de alta qualidade depende significativamente das condições sob as quais o sistema é implementado. É crucial estabelecer essas condições para garantir o desempenho adequado do sistema. A maioria dos sistemas modernos de reconhecimento facial consegue operar eficientemente sob condições específicas. Por exemplo, é essencial organizar o fluxo de pessoas em pontos de controle para permitir a captura facial eficaz a curto prazo. Além disso, a posição das câmeras em relação ao rosto não deve variar mais do que 30 graus. O cumprimento rigoroso dessas condições é fundamental para alcançar a identificação precisa e a busca eficaz de indivíduos, conforme as altas taxas de precisão declaradas pelos fabricantes desses sistemas.

Recentemente, tem havido uma tendência em direção ao uso de software de reconhecimento facial que emprega modelos 3D para aumentar a precisão. O reconhecimento facial 3D utiliza características faciais específicas, como os contornos das cavidades oculares, nariz e queixo, para capturar uma imagem tridimensional em tempo real do rosto de uma pessoa (Figura 2). Essas características são únicas para cada indivíduo e permanecem consistentes ao longo do tempo (Engelbrecht, 2007).

Figura 2 – Modelo de Face tridimensional



Fonte: Elaborado pelos autores.

Utilizando um eixo de profundidade e medição que não é afetado pela iluminação, o reconhecimento facial 3D pode ser empregado mesmo em ambientes com pouca luz, permitindo o reconhecimento do objeto a partir de diferentes ângulos de visualização, incluindo até 90 graus (perfil facial). Com o uso de software com capacidades de reconhecimento 3D, o sistema executa uma série de etapas para verificar a identidade de uma pessoa:

- **Detecção** - a imagem pode ser obtida digitalizando uma fotografia existente (2D) digitalmente ou capturando uma imagem ao vivo do objeto por meio de vídeo (3D);
- **Alinhamento** - após detectar a pessoa, o sistema determina a posição, tamanho e postura da cabeça. No caso de reconhecimento 3D, o sistema é capaz de identificar o objeto em um ângulo de até 90 graus, enquanto em modelos 2D a cabeça deve estar virada para a câmera por pelo menos 35 graus;
- **Mensurações** - o sistema realiza cálculos das curvas da pessoa com uma precisão submilimétrica ou de micro-ondas, criando assim um modelo detalhado;
- **Codificação** - o sistema traduz o modelo em um código único. Esta codificação atribui a cada modelo um conjunto de números que representam as características faciais do objeto;
- **Correspondência** - se a imagem for tridimensional e o banco de dados contiver imagens tridimensionais, a comparação é realizada sem a necessidade de alterações na imagem (Hart, 1968; Wilson, 1972; Tomek, 1976; Kibbler; Aha, 1987; Jankowski, 2000; Reinartz, 2002; Jankowski; Grochowski, 2004; Engelbrecht, 2007).

Atualmente, há um desafio significativo com a autenticação de imagens que permanecem em formato 2D. Em contraste, no contexto do reconhecimento 3D, que envolve a representação de um objeto em três dimensões espaciais, geralmente através das coordenadas X, Y e Z, pontos distintos são determinados, como a parte externa do olho, a parte interna do olho e a ponta do nariz. Esses pontos são marcados e medidos para criar uma imagem 3D. Posteriormente, um algoritmo é aplicado para projetar essa imagem em um formato 2D. Após essa conversão, o programa compara a imagem resultante com as imagens 2D armazenadas no banco de dados para identificar possíveis correspondências.

Existem dois modos principais de operação: verificação e identificação. Na verificação (1:1), a imagem é comparada exclusivamente com uma figura específica no banco de dados para validar a identidade. Em contrapartida, na identificação (1:2), a imagem é comparada com múltiplas figuras no banco de dados, gerando uma pontuação para cada possível correspondência. Esse método é utilizado quando é necessário identificar uma pessoa entre várias.

O modelo vetorial é reduzido e utilizado principalmente para buscas rápidas em bancos de dados, especialmente em buscas de um-para-muitos. A Análise de Textura de Superfície (ATS) é a maior das três metodologias e realiza uma etapa final após a busca no modelo LFA, baseando-se nos elementos da pele na imagem que contêm informações mais detalhadas. Devido à combinação de todos esses padrões, o *Face ID* apresenta uma vantagem sobre outros sistemas. Ele é insensível a mudanças na expressão facial, incluindo piscar, franzir a testa ou sorrir, e pode compensar a presença de barba, bigode e o uso de óculos. No entanto, o *Face ID* não é perfeito. Há diversos fatores que podem impedir o reconhecimento, como reflexos significativos nos óculos, a presença de óculos escuros, cabelo cobrindo a parte central do rosto, iluminação inadequada que resulte em uma imagem mal iluminada e baixa resolução (imagem capturada de muito longe).

No entanto, os fabricantes estão se esforçando para melhorar a usabilidade e a precisão dos sistemas. Os sistemas de reconhecimento facial são utilizados em várias áreas:

- Desbloqueio de telefones: Muitos celulares, incluindo os mais recentes iPhones, utilizam o reconhecimento facial para desbloquear o dispositivo.
- Aplicação da lei: O reconhecimento facial é regularmente usado por agências de segurança.

- Aeroportos e controle de fronteiras: A tecnologia tornou-se comum em muitos aeroportos ao redor do mundo.
- Busca por pessoas desaparecidas: Pode ser utilizado para localizar pessoas desaparecidas e vítimas de tráfico humano.
- Redução de crimes no varejo: É empregado para identificar quando ladrões conhecidos, criminosos organizados do varejo ou pessoas com histórico de fraude entram nas lojas.
- Aprimoramento da experiência de compra: A tecnologia tem o potencial de melhorar a experiência do cliente no varejo.
- Bancos: O banco on-line biométrico é outro benefício do reconhecimento facial.
- *Marketing* e publicidade: Profissionais de *marketing* utilizam a tecnologia para melhorar a experiência do consumidor.
- Saúde: Hospitais utilizam o reconhecimento facial para ajudar pacientes.
- Controle de presença de alunos ou funcionários: Algumas instituições educacionais usam a tecnologia para garantir a presença dos alunos nas aulas. (Hart, 1968; Reinartz, 2002; Jankowski; Grochowski, 2004; Engelbrecht, 2007; Subbotin, 2013a, 2013b).

Além dessas aplicações, o software de gestão de câmeras IP para trabalhadores de PC (NVR) é suportado por um sistema autêntico de controle central, que é uma solução de monitoramento e controle que suporta um número ilimitado de câmeras. O console principal é o servidor de gravação do NVR, que exibe vídeo ao vivo e configura o sistema. O recurso *Save Video* converte imagens para formatos de vídeo padrão. Para garantir brilho, nitidez e tons de cinza uniformes na imagem, é necessário utilizar um aprimorador de vídeo. Outro uso importante da análise de vídeo em sistemas de vigilância é o uso de câmeras inteligentes. Uma câmera inteligente é uma câmera equipada com um módulo adicional que realiza o processamento de vídeo, e esses dois elementos são geralmente integrados em um único corpo.

A principal diferença entre uma câmera inteligente e uma câmera comum é que a câmera inteligente analisa o que vê e toma decisões com base nos resultados dessa análise. Um programa de computador realiza uma análise matemática dos dados de entrada e encontra padrões que podem ser descritos matematicamente. As características faciais únicas de uma pessoa são codificadas em um arquivo de computador usando apenas uma pequena quantidade de memória (menos de 100 bytes). Esse rosto é então comparado com rostos previamente capturados e armazenados em um banco de dados. O operador que trabalha com os dados deve ser informado exibindo informações adicionais sobre o fluxo de vídeo atual. As desvantagens

típicas de sistemas desse tipo incluem funcionalidade limitada, impossibilidade de expansão de *software* e *hardware*, e processamento de dados em "nuvens" corporativas (Subbotin, 2013a, 2013b).

Atualmente, a empresa investe em pesquisa e desenvolvimento para expandir suas capacidades de negócios no sistema de controle de acesso e alarme de segurança. As vantagens desse sistema incluem: Vídeo em 1080p; Notificações instantâneas de atividades; Áudio bidirecional; Imagem perfeita a uma distância de 500 metros, sem perda de qualidade; Funcionamento no escuro e Possibilidade de definir zonas de atividade (Subbotin, 2013a, 2013b). A partir da pesquisa, foi elaborada uma tabela com os requisitos funcionais (Tabela 1).

Tabela 1 – Requisitos funcionais do programa

1.	*O programa deve carregar a base de dados facial*
2.	O programa deve extrair as características faciais das fotos no banco de dados
3.	O programa deve enviar a imagem da *webcam* para o monitor
4.	O programa deve exibir a assinatura do rosto reconhecido ou a palavra "Desconhecido" se o rosto não for reconhecido

Fonte: Elaborado pelos autores.

A escolha da categoria e do método depende das restrições e condições de reconhecimento de pessoas. Entre as restrições que influenciam a escolha do princípio para resolver o problema, destacam-se: a presença ou ausência de obstáculos artificiais no rosto, características espaciais da posição das pessoas, cor da imagem, escala do rosto e resolução da imagem, número de pessoas na imagem, condições de iluminação dos objetos, e a prioridade em minimizar falsos reconhecimentos ou em maximizar a quantidade de pessoas reconhecidas.

Existem diversos métodos e abordagens utilizados em sistemas de reconhecimento facial. Entre eles, destaca-se o método dos Componentes Principais (PCA), a Análise Discriminante Linear (LDA), Modelos Ocultos de Markov (HMM), e *Wavelets* de Gabor (Barina, 2011; Yoon, 2009; Jurafsky; Martin, 2016). Ao utilizar Modelos Ocultos de Markov para resolver o problema de reconhecimento facial, cada classe de rostos calcula seu próprio modelo oculto de Markov. Em seguida, para o método desconhecido, todos os modelos disponíveis são executados, e busca-se aquele que proporciona o resultado mais próximo. A desvantagem dessa abordagem é que os Modelos Ocultos de Markov não possuem boa resolução, pois o algoritmo de aprendizado maximiza a resposta para suas classes, mas não minimiza a resposta para outras classes.

Métodos de reconhecimento baseados no uso de Wavelets de Gabor mostram alta eficiência. Filtros de Gabor são utilizados na etapa de pré-processamento para formar um vetor de características de Gabor de uma imagem facial. O método das Wavelets de Gabor é resistente a mudanças na iluminação, pois não utiliza diretamente os valores dos tons de cinza de cada pixel, mas extrai características (Hart, 1968; Aha; Kibler; Albert, 1991; Reinartz, 2002; Jankowski; Grochowski, 2004).

A seguir, o trabalho descreve e analisa métodos modernos de reconhecimento facial, começando pelo Método dos Componentes Principais (PCA). A ideia é representar imagens faciais como um conjunto de componentes principais das imagens, conhecidos como "rostos próprios". Estes rostos possuem a propriedade útil de que cada vetor correspondente se assemelha ao formato de um rosto. O cálculo dos componentes principais é realizado através da obtenção dos autovetores e autovalores da matriz de covariância derivada da imagem. A reconstrução da imagem é obtida pela combinação linear dos principais componentes multiplicados pelos vetores próprios correspondentes (Hart, 1968; Jankowski; Grochowski, 2004).

Para cada fotografia facial, são calculados seus principais componentes, variando geralmente de 5 a 200. O processo de reconhecimento envolve a comparação dos principais componentes do retrato desconhecido com os componentes de todas as imagens conhecidas. Supõe-se que as imagens correspondentes à mesma pessoa formem clusters no espaço de características. São selecionadas as fotos que apresentam menor distância em relação à imagem de entrada (desconhecida) no banco de dados (Jankowski; Grochowski, 2004).

O método dos rostos próprios requer condições ideais para sua aplicação, como iluminação uniforme, expressão facial neutra e ausência de obstruções como óculos e barba. Se essas condições não forem atendidas, os componentes principais podem não captar variações entre indivíduos. Em diferentes condições de iluminação, este método pode ser limitado, pois os primeiros componentes principais frequentemente refletem variações na iluminação, resultando em comparações que não são discriminativas. Sob condições ideais, a precisão de reconhecimento utilizando este método pode alcançar mais de 90%, representando um resultado significativo

O cálculo de um conjunto de autovetores é altamente demorado. Um dos métodos para mitigar isso é reduzir as imagens por linhas e colunas, resultando em uma representação menor que acelera o processo de cálculo e reconhecimento, embora não seja mais possível restaurar a foto original. O Método Viola-Jones é altamente eficaz na detecção de objetos em imagens e

vídeos em tempo real. Esta abordagem apresenta uma baixa probabilidade de identificação errônea de uma pessoa, sendo capaz de detectar características faciais mesmo sob pequenos ângulos, aproximadamente até 30 graus. A precisão de reconhecimento utilizando este método pode superar 90% (Winarno *et al.,* 2018).

O Método de Comparação de Modelos (*Template Matching*) baseia-se na seleção e comparação de áreas específicas do rosto em diferentes imagens. Cada área similar aumenta o grau de similaridade entre as fotografias. Algoritmos simples, como comparação de *pixels*, são utilizados para esta conferência (Hart, 1968; Reinartz, 2002; Jankowski; Grochowski, 2004). Uma desvantagem deste método é a demanda por recursos significativos tanto para armazenamento quanto para comparação de áreas. Além disso, requer que as imagens sejam capturadas sob condições estritamente controladas, sem variações perceptíveis na pose, iluminação, expressão facial, entre outros. A precisão de reconhecimento utilizando este método é em torno de 80% (Jankowski; Grochowski, 2004).

A Rede Neural Hopfield utiliza um algoritmo de aprendizagem que difere dos métodos clássicos de perceptron, calculando todos os coeficientes da matriz de pesos em um único ciclo. Esta rede é configurada rapidamente para o reconhecimento de padrões. No entanto, suas limitações incluem a sensibilidade a imagens muito semelhantes e a necessidade de que as imagens não sejam deslocadas ou rotacionadas em relação ao estado original. Para superar estas limitações, diversas modificações da rede neural Hopfield têm sido propostas, incluindo transformações ortogonais que permitem a recuperação de imagens altamente correlacionadas através de conjuntos duais de vetores. A precisão de reconhecimento utilizando este método pode exceder 90%, alcançando em alguns casos resultados próximos a 100% (Hart, 1968; Aha; Kibler; Albert, 1991; Broadley, 1993; Jankowski, 2000; Wilson, 2000; Reinartz, 2002; Jankowski; Grochowski, 2004).

Uma abordagem abrangente para o reconhecimento facial, que se mostrou promissora na resolução da tarefa proposta, envolve a implementação de experimentos computacionais que utilizam imagens de meio-tom através de classificadores Euclidiano-Mahalanobis (baseados em métricas), redes neurais probabilísticas e momentos invariantes. O uso da métrica Euclidiano-Mahalanobis permite ao sistema lidar com variações na orientação da cabeça e mudanças na luminosidade da imagem. Por outro lado, as redes neurais probabilísticas são eficazes na gestão de desafios como olhos fechados e variações na expressão facial (como sorrisos e caretas). Entretanto, o treinamento de inteligência artificial para redes neurais

demanda um investimento significativo de tempo devido à quantidade extensa de dados necessária.

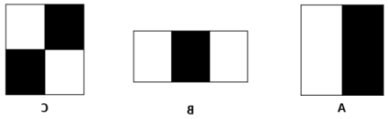
Deve-se destacar uma vantagem significativa da tecnologia de reconhecimento facial: sua inviolabilidade, uma vez que não envolve elementos como senhas que possam ser roubados ou alterados. Não é viável garantir que uma pessoa mantenha uma posição fixa para permanecer parada e olhar diretamente para a câmera, o que ocasionalmente compromete a precisão dos resultados. Se uma pessoa modificar sua aparência, como mudar o cabelo ou usar acessórios, pode se tornar praticamente impossível reconhecê-la (Broadley, 1993; Jankowski, 2000; Wilson, 2000; Reinartz, 2002; Jankowski; Grochowski, 2004).

Com base nas considerações anteriores, a criação de métodos híbridos que combinem as vantagens e minimizem as desvantagens das abordagens individuais mencionadas parece ser uma direção promissora para o avanço contínuo da tecnologia de reconhecimento facial.

Um algoritmo de particular destaque é o Viola-Jones. Embora seja lento para o aprendizado, este algoritmo é capaz de detectar faces em tempo real com uma eficiência impressionante. Funcionando em imagens em escala de cinza, o algoritmo examina várias sub-regiões menores da imagem e procura por características específicas em cada uma delas para identificar rostos. É necessário verificar muitas posições e escalas diferentes, pois uma imagem pode conter múltiplas faces de diferentes tamanhos. Viola e Jones utilizaram funções semelhantes a Haar para a detecção de rostos neste algoritmo (Reinartz, 2002; Jankowski; Grochowski, 2004).

É importante notar que este método apresenta uma probabilidade muito baixa de identificação falsa de uma pessoa. No entanto, a eficácia da detecção diminui significativamente em ângulos superiores a 30 graus, dificultando a detecção de rostos humanos posicionados em ângulos variáveis. Esse aspecto pode complicar a implementação do algoritmo em sistemas de produção modernos, considerando as crescentes demandas por precisão. O algoritmo se baseia em três tipos de características semelhantes a Haar, conforme ilustrado na Figura 3.

Figura 3 – Três Tipos de Características Viola-Jones

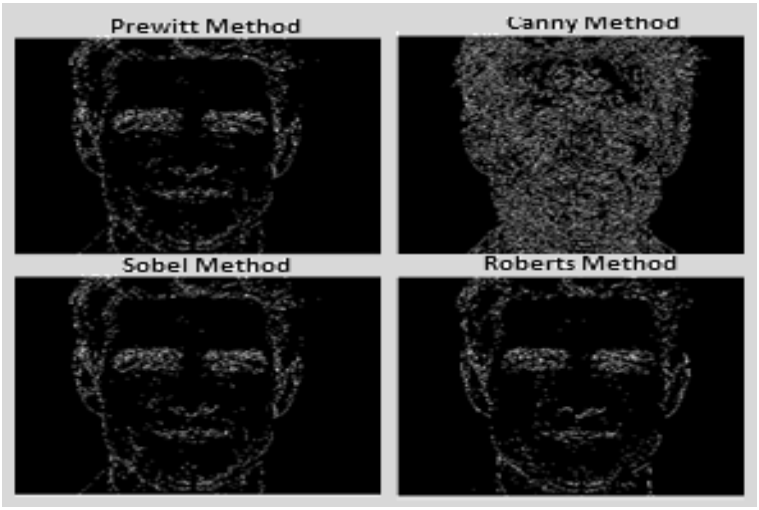


Fonte: Elaborado pelos autores.

A representação integral da imagem é uma matriz que coincide em tamanho com a foto original. Cada elemento armazena a soma das intensidades de todos os pixels localizados à esquerda e acima do elemento dado. A *Wavelet transform* (WT) é frequentemente empregada para analisar processos instáveis, demonstrando eficácia na aplicação a diversas tarefas relacionadas ao processamento de imagens. Os coeficientes contêm informações sobre o processo analisado e a *Wavelet* utilizada, sendo a escolha desta, determinada pelo tipo de informação a ser extraída do processo. Cada *Wavelet* possui características distintas durante o movimento linear nos domínios temporal e de frequência.

Para a determinação dos contornos faciais, foram conduzidos diversos experimentos utilizando o MatLab, um sistema de matemática computacional que utiliza as bibliotecas de Visão Computacional (CV) e *Toolbox*. Os resultados dos experimentos indicaram que o método de *Prewitt* e *Sobel* foi o mais eficaz dentro dos limites de aplicação especificados (Fig. 4). A taxa de detecção facial foi de 75%, com uma variação de inclinação entre 14% e 26%. Estes experimentos destacaram que os métodos de *Sobel* e *Prewitt* são os mais eficientes para a seleção de contornos, devido à sua capacidade de refletir com precisão os principais e secundários contornos da imagem sem introduzir ruído excessivo. Tais métodos serão adotados em futuras aplicações.

Figura 4 – Métodos de Seleção de Contornos



Fonte: Elaborado pelos autores.

Para o sistema de Reconhecimento Facial, foi adotada a linguagem de programação Python, e dentro dessa linguagem, a biblioteca *OpenCV* é utilizada. *OpenCV* é uma biblioteca de visão computacional de código aberto projetada para análise, classificação e processamento

de imagens. Amplamente empregada em linguagens como C, C++, Python e Java, essa biblioteca é aplicada na detecção de objetos, rostos, doenças, lesões, placas de veículos e até mesmo texto manuscrito em diversas imagens e vídeos. Com o *OpenCV* em Aprendizado Profundo (*Deep Learning*), expande-se o espaço vetorial e realizam-se operações matemáticas sobre essas características para identificar padrões visuais e suas diversas características.

A visão computacional é definida como uma disciplina que explora como reconstruir, extrair e entender um espaço 3D a partir de suas imagens 2D, em termos das propriedades da estrutura presente no espaço. Ela dedica-se à modelagem e à replicação da visão humana utilizando *software* e *hardware* de computador (Wilson, 1972).

A visão computacional é o processo pelo qual se compreende imagens e vídeos, determinando como são armazenados e como podem ser manipulados e extraídos dados deles. Essa tecnologia é fundamental para a inteligência artificial, desempenhando um papel importante em carros autônomos, robótica e aplicativos de correção de fotos. Diferentemente da fotografia, a visão computacional não reflete a realidade de forma direta, mas a interpreta, muitas vezes incorretamente, impondo significados baseados em suposições estatísticas. Os dados brutos dos algoritmos de visão computacional não possuem valor verdadeiro intrínseco; eles fornecem apenas probabilidades estatísticas que são truncadas em estados booleanos, disfarçados de resultados verdadeiros, com significados adicionados na pós-produção.

Dlib é uma das bibliotecas de código aberto mais poderosas e fáceis de usar, consistindo em bibliotecas e algoritmos de aprendizado de máquina, além de várias ferramentas de desenvolvimento de software. A licença de código aberto do Dlib permite seu uso gratuito em qualquer aplicação (Tomek, 1976).

O Dlib oferece duas funções distintas para captura de rosto: HoG + *Linear SVM* e *Max-Margin CNN*. O algoritmo *Histogram Oriented Gradients* (HoG) + *Linear Support Vector Machine* (SVM) no Dlib proporciona reconhecimento facial frontal muito rápido, mas possui capacidades limitadas para reconhecer poses faciais em ângulos agudos, como em imagens de CCTV ou em ambientes de observação aleatória, onde o sujeito não participa ativamente do processo de identificação. Este método é adequado para situações onde o sensor pode contar com uma visão direta e desobstruída do rosto do participante, como em caixas eletrônicos, sistemas de identificação de moldura móvel e sistemas de reconhecimento CCTV móvel, onde as câmeras podem obter um perfil direto de motoristas.

O detector de rosto *Max-Margin CNN* (MMOD) é um detector robusto e confiável, acelerado por GPU, que utiliza uma Rede Neural Convolucional (CNN). Este método é

significativamente melhor na detecção de rostos em ângulos obscuros e em ambientes desafiadores, sendo adequado para vigilância casual e análise urbana.

O MMOD (*Maximum-Margin Object Detection*) não é uma alternativa independente ao HoG (*Histogram of Oriented Gradients*) + Linear SVM, mas pode ser aplicado ao próprio HoG ou a qualquer modelo visual de saco de palavras que trate clusters de pixels detectados como entidades para possível rotulagem, incluindo o reconhecimento facial. A seguir, serão comparados o HoG e o MMOD. A atratividade do HoG + Linear SVM no Dlib está em seu baixo uso de recursos, sua eficácia ao trabalhar na CPU, e sua capacidade de lidar com rostos não frontais, além de oferecer um procedimento relativamente poderoso de detecção de oclusão. No entanto, a implementação padrão exige um tamanho mínimo de rosto de 80×80 *pixels*. Para detectar rostos menores que esse limite, é necessário treinar uma implementação própria. Além disso, essa abordagem apresenta resultados insatisfatórios para ângulos agudos do rosto, cria quadros de delimitação que podem cortar excessivamente as características faciais e enfrenta dificuldades em casos complexos de oclusão.

Por outro lado, a vantagem do MMOD (CNN) no Dlib reside principalmente em sua capacidade de reconhecer orientações faciais complexas, uma característica decisiva dependendo do ambiente-alvo. Ele também oferece uma velocidade impressionante com acesso a uma GPU média, uma arquitetura de aprendizado fácil e excelente processamento de oclusão. No entanto, ele pode produzir retângulos de delimitação ainda mais restritos que o HoG + Linear SVM na implementação padrão, opera significativamente mais lento na CPU do que o HoG/LSVM, e compartilha a incapacidade do HoG/LSVM de detectar rostos menores que 80 *pixels*, novamente exigindo uma construção personalizada para certos cenários, como pontos de vista de ruas que se estendem à distância.

A principal tarefa do sistema de "Reconhecimento Facial" é a detecção de rostos. Primeiro, é necessário encontrar um rosto em uma foto ou imagem antes de poder distingui-los. Quando a câmera detecta rostos automaticamente, ela entende que todos os rostos estão em foco antes de tirar a foto. Nesse caso, tal função é utilizada para detectar uma parte da imagem que será posteriormente usada para reconhecimento.

Neste trabalho, foi escolhido o método "Histograma de Gradientes Orientados" (HoG) por sua confiabilidade. Para encontrar um rosto em uma imagem, é necessário torná-la preta e branca, pois os dados de cor não são necessários para este processo. Para converter uma foto ou imagem em preto e branco, utiliza-se uma função da biblioteca OpenCV (Fig. 5).

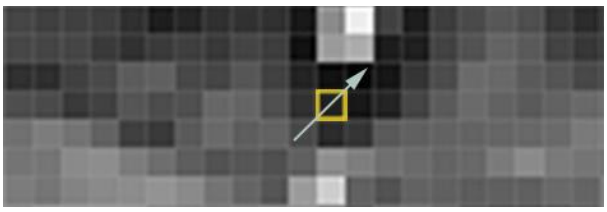
Figura 5 – Convertendo a Imagem para Preto e Branco

```
def findEncodings(images):
    encodeList = []
    for img in images:
        img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
```

Fonte: Elaborado pelos autores.

Cada pixel da imagem é então examinado individualmente. Nesse processo, é necessário considerar todos os *pixels* ao redor de cada *pixel* isolado. A principal tarefa deste procedimento é identificar o *pixel* mais escuro da imagem e coletar dados sobre os outros *pixels* ao seu redor. O próximo passo será determinar a direção em que a imagem se torna mais escura, representada por setas (Fig. 6).

Figura 6 – Gradientes de *Pixels*



Fonte: Elaborado pelos autores.

Repetindo este procedimento para cada pixel da imagem, todos os *pixels* serão substituídos por setas. Essas setas, chamadas de gradientes, indicam o fluxo dos *pixels* claros para os escuros ao longo da imagem. A substituição dos *pixels* por gradientes é vantajosa na análise comparativa de imagens, pois tanto imagens escuras quanto claras da mesma pessoa terão valores de *pixels* diferentes. Contudo, ao considerar a direção da mudança no brilho da foto, as imagens claras e escuras resultarão na mesma representação. A desvantagem desta abordagem é que, como resultado, obteremos um grande número de gradientes, o que dificultará o reconhecimento facial posterior.

Para mitigar essa complexidade, a imagem é dividida em quadrados de 256 pixels, ou seja, quadrados com lados de 16 por 16. O próximo passo é calcular a quantidade de gradientes que indicam a direção do brilho. Contabiliza-se quantos gradientes apontam para cima, para o canto superior direito, para a direita e assim por diante. Nesse quadrado, o gradiente é substituído pelas direções das setas que predominam. Como resultado, obtém-se uma imagem simplificada na qual a estrutura principal do rosto é visível. Para encontrar um rosto em uma imagem HOG, basta localizar a parte de outra imagem que seja mais semelhante ao padrão HOG derivado de vários rostos de treinamento (Figura 7).

Figura 7 – Comparação de Características

Fonte: Elaborado pelos autores.

Para cumprir essas condições, utiliza-se a função de localização de rostos da biblioteca "*face_recognition*", que emprega o método de detecção (Figura 8). A tecnologia facilita a identificação de rostos em qualquer imagem.

Figura 8 – Detecção Facial

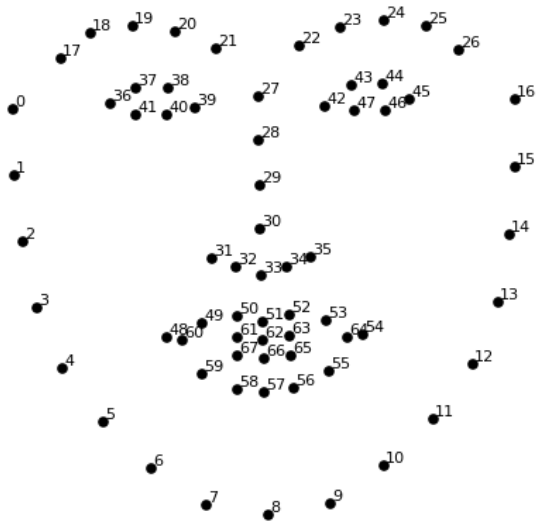
```
faceLoc = face_recognition.face_locations(imgRus)[0]
encodeElon = face_recognition.face_encodings(imgRus)[0]
cv2.rectangle(imgRus, (faceLoc[3], faceLoc[0]),
              (faceLoc[1], faceLoc[2]), (255, 0, 255), 2)

faceLocTest = face_recognition.face_locations(imgTest)[0]
encodeTest = face_recognition.face_encodings(imgTest)[0]
cv2.rectangle(imgTest,(faceLocTest[3],faceLocTest[0]),
              (faceLocTest[1],faceLocTest[2]),(255,0,255),2)
```

Fonte: Elaborado pelos autores.

A próxima tarefa é a formação de pontos de referência faciais. O rosto na imagem é isolado. A partir desse ponto, surge um problema: o rosto humano pode estar virado em diferentes direções e, por isso, parecer diferente para a IA. Para resolver esse problema, é necessário garantir que os lábios e os olhos estejam alinhados na imagem da mesma forma que no modelo. Isso simplificará significativamente a comparação de rostos nas etapas subsequentes. Para isso, é utilizado um algoritmo chamado "estimação de pontos de referência facial". Existem muitos métodos para resolver esse problema, mas nesta situação será utilizado o método de Kazemi e Sullivan. A ideia básica é destacar 68 pontos específicos (chamados *landmarks*) que existem em todo rosto – o topo do queixo, a borda externa de cada olho, a borda interna de cada sobrancelha, e assim por diante. Depois disso, o algoritmo é programado para encontrar esses 68 pontos específicos em diferentes formas e tipos de rosto (Figura 9).

Neste ponto, após o algoritmo ser treinado para reconhecer rostos e bocas, resta apenas rotacionar, dimensionar e transladar a imagem de modo que os olhos e a boca estejam o mais centralizados possível. Em uma transformação afim, todas as linhas paralelas na imagem original permanecerão paralelas na imagem resultante.

Figura 9 – Pontos-chave do rosto (*Landmarks*)



Fonte: Elaborado pelos autores.

Para encontrar a matriz de transformação, são necessários três pontos da imagem de entrada e suas respectivas localizações na imagem de saída. Uma transformação afim utiliza um ângulo de rotação no sentido horário, ao contrário do círculo unitário da geometria típica, onde a rotação é medida no sentido anti-horário a partir do 0, começando no eixo X positivo. Por isso, frequentemente se vê um valor de ângulo negativo sendo utilizado.

Uma transformação afim é caracterizada pelas seguintes propriedades: qualquer transformação afim pode ser representada como uma sequência de operações simples, como translação, estiramento/compressão e rotação. Linhas retas, o paralelismo das linhas retas, a razão das distâncias de segmentos situados na mesma linha e a razão das áreas das figuras são preservados. As novas coordenadas f(x), que possuem a mesma posição no espaço relativo ao "novo" sistema de coordenadas, que as coordenadas x tinham no "antigo". Portanto, será suficiente utilizar as transformações básicas: escala e rotação, preservando as linhas paralelas, ou seja, laços afins.

Assim, completa-se uma das principais tarefas: a centralização dos olhos e da boca. Independentemente do ângulo do rosto na imagem, o algoritmo centralizará esses pontos na mesma posição, o que é crucial para a precisão das etapas subsequentes.

Para identificar uma pessoa, utilizamos a abordagem de comparar um rosto desconhecido com as características já descritas em uma base de dados de fotos coletadas. É necessário que o rosto desconhecido não apenas se pareça com uma pessoa já registrada na base de dados, mas que seja, de fato, essa pessoa. Esta abordagem apresenta vários desafios. Por exemplo, um site como o Facebook, com bilhões de usuários e trilhões de fotos, não pode comparar todos os rostos previamente marcados com cada nova imagem carregada, pois isso demandaria um tempo excessivo.

Em vez disso, os algoritmos precisam ser capazes de reconhecer rostos em milissegundos, não em horas. Para evitar sobrecarregar o dispositivo durante a comparação, é necessário selecionar os principais parâmetros de algumas partes do rosto para serem comparados por meio de cálculos. Para determinar a abordagem mais precisa para a comparação, foram conduzidos uma série de experimentos que revelaram que, nesta situação, é mais eficiente permitir que o computador escolha independentemente quais medições coletar. Nesse aspecto, o algoritmo executa a tarefa de maneira mais eficaz do que um humano.

A solução envolve o treinamento de uma rede neural convolucional profunda. Essa rede é treinada para gerar 128 valores únicos para qualquer rosto, o que é mais eficaz do que treiná-la para reconhecer rostos diretamente. A capacitação da rede envolve a visualização simultânea de três tipos de imagens faciais: uma imagem de treino do rosto de uma pessoa famosa, outra imagem do mesmo rosto e uma imagem de uma pessoa completamente diferente. O algoritmo, então, compara os valores gerados para as três imagens.

O próximo passo é ajustar a rede neural para garantir que as medições geradas para as duas imagens da mesma pessoa sejam mais semelhantes entre si do que em relação à imagem da pessoa diferente. Repetindo essa etapa milhões de vezes para imagens de milhares de pessoas diferentes, a rede neural aprende a gerar 128 parâmetros únicos para cada rosto. Dez fotos diferentes da mesma pessoa devem resultar em medições semelhantes. No campo do aprendizado de máquina, esses 128 parâmetros de cada rosto são chamados de "*embedding*". Codificar uma imagem facial envolve treinar uma rede neural convolucional para extrair essas *embeddings*, requerendo uma quantidade significativa de dados e poder computacional. Mesmo com uma placa de vídeo Nvidia avançada, como a Tesla, são necessárias cerca de 24 horas de treinamento contínuo para alcançar uma boa precisão (Jankowski, 2000).

Após o treinamento, a rede neural adquire a capacidade de gerar medições para todos os tipos de rostos, mesmo que nunca os tenha visto antes. Portanto, essa etapa de treinamento precisa ser realizada apenas uma vez. Graças aos esforços dos colaboradores do projeto *OpenFace*, que foram treinados e capacitados, seu trabalho resultou na publicação de várias redes pré-treinadas, uma das quais foi utilizada neste estudo. Posteriormente, as imagens faciais são processadas pela rede pré-treinada para obter 128 medições únicas para cada rosto (Figura 10).

Figura 10 – Codificação Facial

Fonte: Elaborado pelos autores.

A rede neural gera os mesmos valores numéricos ao comparar duas imagens diferentes da mesma pessoa. Para utilizar uma rede neural pré-treinada em equipamentos potentes, emprega-se a função de codificação facial da biblioteca *Face Recognition* (Figura 11).

Figura 11 – Codificação Facial

```
faceLoc = face_recognition.face_locations(imgRus)[0]
encodeElon = face_recognition.face_encodings(imgRus)[0]
cv2.rectangle(imgRus, (faceLoc[3], faceLoc[0]),
              (faceLoc[1], faceLoc[2]), (255, 0, 255), 2)

faceLocTest = face_recognition.face_locations(imgTest)[0]
encodeTest = face_recognition.face_encodings(imgTest)[0]
cv2.rectangle(imgTest,(faceLocTest[3],faceLocTest[0]),
              (faceLocTest[1],faceLocTest[2]),(255,0,255),2)
```

Fonte: Elaborado pelos autores.

A última etapa envolve o algoritmo identificar a pessoa no banco de dados de rostos conhecidos cuja medição mais se aproxima da imagem de teste. Para isso, é necessário preparar um classificador que analise a imagem de teste e indique a pessoa correspondente. Esse classificador realiza a análise em milissegundos, resultando na identificação do nome da pessoa.

Todas as funções mencionadas anteriormente são então utilizadas para obter o resultado final (Figura 12).

Figura 12 – Reconhecimento Facial

```
while True:
    success, img = cap.read()
    imgS = cv2.resize(img,(0,0),None,0.25,0.25)
    imgS = cv2.cvtColor(imgS, cv2.COLOR_BGR2RGB)
    facesCurFrame = face_recognition.face_locations(imgS)
    encodesCurFrame = face_recognition.face_encodings(imgS,facesCurFrame)
    for encodeFace,faceLoc in zip(encodesCurFrame,facesCurFrame):
        matches = face_recognition.compare_faces(encodeListKnown,encodeFace)
        faceDis = face_recognition.face_distance(encodeListKnown,encodeFace)
        matchIndex = np.argmin(faceDis)
        if faceDis[matchIndex] < 0.50:
            name = classNames[matchIndex].upper()
            markAttendance(name)
        else:
            name = 'Unknown'
        y1, x2, y2, x1 = faceLoc
        y1, x2, y2, x1 = y1 * 4, x2 * 4, y2 * 4, x1 * 4
        cv2.rectangle(img, (x1, y1), (x2, y2), (0, 255, 0), 2)
        cv2.rectangle(img, (x1, y2 - 35), (x2, y2), (0, 255, 0), cv2.FILLED)
        cv2.putText(img, name, (x1 + 6, y2 - 6), cv2.FONT_HERSHEY_COMPLEX, 1, (255, 255, 255), 2)
    cv2.imshow('Webcam',img)
    cv2.waitKey(1)
```

Fonte: Elaborado pelos autores.

Nesta etapa, é possível testar o programa e determinar a precisão com que a IA reconhece rostos. Para este experimento, foram utilizadas várias imagens de pessoas famosas. Para verificar a funcionalidade do software desenvolvido, foram realizados testes do sistema após a conclusão dos testes de módulo e de integração. Em seguida, o software foi testado no ambiente esperado. Foram empregados métodos de teste funcional e alguns métodos de teste estrutural. O teste de sistema assegura que cada função do sistema esteja operando corretamente e também verifica requisitos não funcionais, como desempenho, segurança, confiabilidade, estresse e carga. Isso, no futuro, proporcionará a oportunidade de melhorar a qualidade do produto final.

Uma análise dos defeitos encontrados na fase de teste de sistema foi realizada. Antes de eliminar qualquer defeito, foi feita uma análise de seu impacto. Caso o sistema permita, os defeitos são simplesmente documentados e mencionados como limitações conhecidas, ao invés de serem corrigidos, se a correção for demorada ou tecnicamente impossível no design atual (Jankowski, 2000).

A lista de condições sob as quais o teste ocorrerá inclui: determinação da pessoa cadastrada no banco de dados; identificação de uma pessoa não cadastrada no banco de dados; identificação de uma pessoa em diferentes condições de iluminação; identificação do rosto em diferentes ângulos; identificação de uma pessoa com diferentes expressões faciais; e identificação de uma pessoa com barba, bigode, óculos ou máscara. Para verificação, foram compilados casos de teste, segundo os quais o programa foi verificado (Tabela 2).

Tabela 2 – Teste para verificação do programa

№	Descrição	Ação	Resultado esperado
1.	Determinação da pessoa inserida no banco de dado.	Aguardar os resultados do reconhecimento.	Pessoa reconhecida (nome exibido).
2.	Identificação de uma pessoa não inserida na base de dados.	Aguardar os resultados do reconhecimento.	A pessoa não é reconhecida.
3.	Identificação de uma pessoa em iluminação adequada (lâmpada diurna de 900Lm).	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
4.	Identificação de uma pessoa no escuro.	Aguardar os resultados do reconhecimento.	A pessoa não é reconhecida.
5.	Detecção de rosto em um ângulo de 16–26%.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
6.	Identificação de rosto em um ângulo de 26–40%.	Aguardar os resultados do reconhecimento.	A pessoa é parcialmente reconhecida.
7.	Identificação de uma pessoa com barba.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
8.	Identificação de uma pessoa com bigode.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
9.	Identificação de pessoa com óculos de lentes transparentes.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
10.	Identificação de uma pessoa com óculos de lentes escuras.	Aguardar os resultados do reconhecimento.	A pessoa não é reconhecida.
11.	Identificação de uma pessoa com máscara médica cobrindo o nariz.	Aguardar os resultados do reconhecimento.	A pessoa não é reconhecida.
12.	Identificação de uma pessoa com máscara médica que não cubra o nariz.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.
13.	Identificação de uma pessoa sorrindo.	Aguardar os resultados do reconhecimento.	A pessoa é reconhecida.

Fonte: Elaborado pelos autores.

Considerações finais

A novidade científica dos resultados obtidos reside no aprimoramento do sistema de reconhecimento facial, especificamente no reconhecimento biométrico por meio do "*Face ID*". Foram investigadas as áreas de atividade em que o sistema de "Reconhecimento Facial" é utilizado e para quais finalidades. Cada componente do sistema foi projetado e desenvolvido progressivamente, culminando na construção do próprio sistema de reconhecimento facial. Para assegurar o correto funcionamento do sistema, foram selecionadas várias fotos com imagens de diferentes pessoas para testes.

Este trabalho incluiu a análise do campo de reconhecimento facial, destacando a relevância atual desse sistema, o reconhecimento biométrico do sistema "*Face ID*", e diversos métodos de reconhecimento facial. Além disso, foi investigado em quais áreas de atividade o sistema de "Reconhecimento Facial" é aplicado e com quais objetivos.

No decorrer do estudo, diversos algoritmos de reconhecimento facial foram analisados. Com base nos resultados dessa análise, comprovou-se que o sistema de reconhecimento facial pode ser modelado utilizando o método de extração de contornos de Viola-Jones, alcançando uma taxa de reconhecimento bem-sucedido de aproximadamente 75%. As capacidades funcionais da biblioteca de visão computacional *OpenCV*, entre outras bibliotecas, foram consideradas.

A importância prática dos resultados obtidos reside no desenvolvimento de um software que implementa os indicadores propostos, e os experimentos conduzidos confirmam a eficácia do desenvolvimento proposto. Os resultados experimentais sustentam a recomendação do produto de *software* para aplicação prática, além de identificar as condições ideais para sua utilização.

Cada componente do sistema foi desenvolvido de maneira progressiva, resultando na construção do próprio sistema de reconhecimento facial. Naturalmente, foi realizada uma verificação para assegurar o funcionamento adequado do sistema, com a seleção de diversas fotos contendo imagens de diferentes pessoas. O plano de teste elaborado reflete as principais etapas da implementação tanto da pesquisa teórica quanto da prática realizada. A solução de *software* proposta pode ser aplicada para implementar sistemas mais robustos, como sistemas de vigilância por vídeo com reconhecimento inteligente.

As perspectivas para pesquisas futuras incluem aprimorar o desenvolvimento proposto para aumentar sua eficiência de uso.

REFERÊNCIAS

AHA, D. W.; KIBLER, D.; ALBERT, M. K. Instance-based learning algorithms. **Machine Learning**, [*S. l.*], v. 6, p. 37–66, 1991. DOI: 10.1023/A:1022689900470.

BARINA, D. Gabor wavelets in image processing. *In*: CONFERENCE STUDENT EEICT, 17., 2011. **Proceedings** [...]. Brno: Brno University of Technology. 2011. v. 3, p. 522–526.

BROADLEY, C. E. Addressing the selective superiority problem: automatic algorithm/model class selection. *In*: MACHINE LEARNING: TENTH INTERNATIONAL CONFERENCE, Amherst, 1993. **Proceedings** […]. Burlington: Morgan Kaufmann, 1993. p. 17–24. DOI: 0.1016/b978-1-55860-307-3.50009-5.

DOMINGO, C.; GAVALDA, R.; WATANABE, O. Adaptive sampling methods for scaling up knowledge discovery algorithms. *In*: DISCOVERY SCIENCE: SECOND INTERNATIONAL CONFERENCE, 1999, Tokyo. **Proceedings** […]. Berlin: Springer, 1999. p. 172–183. DOI: 10.1007/3-540-46846-3_16.

ENGELBRECHT, A. **Computational intelligence**: an introduction. Sidney: John Wiley & Sons, 2007. 597 p. DOI: 10.1002/9780470512517.

EVANS, R. **Clustering for classification**: using standard clustering methods to summarise datasets with minimal loss of classification accuracy. Saarbrücken: VDM Verlag, 2008. 108 p.

GATES, G. The reduced nearest neighbor rule. **IEEE Transactions on Information Theory**, [*S. l.*], v. 18, n. 3, p. 431–433, 1972. DOI: 10.1109/TIT.1972.1054809.

HART, P. E. The condensed nearest neighbor rule. **IEEE Transactions on Information Theory**, [*S. l.*], v. 14, p. 515–516, 1968. DOI: 10.1109/TIT.1968.1054155.

JANKOWSKI, N. Data regularization. *In*: NEURAL NETWORKS AND SOFT COMPUTING, 5., 2000, Zakopane. **Proceedings** […]. Częstochowa: Polish Neural Networks Society, 2000. p. 209–214.

JANKOWSKI, N.; GROCHOWSKI, M. Comparison of instance selection algorithms I. Algorithms survey. *In:* ARTIFICIAL INTELLIGENCE AND SOFT COMPUTING, 7., 2004, Zakopane. **Proceedings** […]. Berlin: Springer, 2004. p. 598–603. (Lecture Notes in Computer Science, v. 3070). DOI: 10.1007/978-3-540-24844-6_90.

JURAFSKY, D.; MARTIN, J. **Hidden Markov models**. Speech and Language Processing. Draft of November 7, 2016.

KIBBLER, D.; AHA, D. W. Learning representative exemplars of concepts: an initial case of study. *In*: MACHINE LEARNING INTERNATIONAL WORKSHOP, 4., 1987, Irvine. **Proceedings** […]. Burlington: Morgan Kaufmann, 1987. p. 24–30. DOI: 10.1016/b978-0-934613-41-5.50006-4.

KOHONEN, T. Learning vector quantization. **Neural Networks**, [*S. l.*], v. 1, p. 303, 1988. DOI: 10.1016/0893-6080(88)90334-6.

KOSKIMAKI, H.; JUUTILAINEN, I.; LAURINEN, P.; RÖNING, J. Two-level clustering approach to training data instance selection: a case study for the steel industry. *In*: NEURAL NETWORKS: INTERNATIONAL JOINT CONFERENCE, 2008, Hong Kong. **Proceedings** […]. Los Alamitos: IEEE, 2008. p. 3044–3049. DOI: 10.1109/ijcnn.2008.4634228.

LI, B.; CHI, M.; FAN, J.; XUE, X. Support cluster machine. *In:* MACHINE LEARNING INTERNATIONAL CONFERENCE, 24., 2007, Corvallis. **Proceedings** […]. New York, 2007. p. 505–512. DOI: 10.1145/1273496.1273560.

MADIGAN, D.; RAGHAVAN, N.; DUMOUCHEL, W.; NASON, M. Likelihood-based data squashing: a modeling approach to instance construction. **Data Mining and Knowledge Discovery**, [*S. l.*], v. 6, n. 2. p. 173–190. DOI: 10.1023/A:1014095614948.

REEVES, C. R.; BUSH, D. R. Using genetic algorithms for training data selection in RBF networks. *In*: **Instance Selection and Construction for Data Mining.** Norwell: Kluwer, 2001. Part VI. p. 339–356. DOI: 10.1007/978-1-4757-3359-4_19.

REINARTZ, T. A. Unifying view on instance selection. **Data Mining and Knowledge Discovery**, [*S. l.*], v. 6, p. 191–210, 2002. DOI: 10.1023/A:1014047731786.

RITTER, G.; WOODRUFF, H.; LOWRY, S.; ISENHOUR, T. An algorithm for a selective nearest neighbor decision rule. **IEEE Transactions on Information Theory**, [*S. l.*], v. 21, n. 6, p. 665–669, 1975. DOI: 10.1109/TIT.1975.1055464.

SANE, S. S.; GHATOL, A. A. A Novel supervised instance selection algorithm. **International Journal of Business Intelligence and Data Mining,** [*S. l.*], v. 2, n 4. p. 471–495, 2007. DOI: 10.1504/IJBIDM.2007.016384.

SKALAK, D. B. Prototype and feature selection by sampling and random mutation hill climbing algorithms. *In*: MACHINE LEARNING: INTERNATIONAL CONFERENCE, 11., 1994, New Brunswick. **Proceedings** […]. Burlington: Morgan Kaufmann, 1994. p. 293–301. DOI: 10.1016/b978-1-55860-335-6.50043-x.

SUBBOTIN, S. The neuro-fuzzy network synthesis and simplification on precedents in problems of diagnosis and pattern recognition. **Optical Memory and Neural Networks (Information Optics)**, [*S. l.*], v. 22, n. 2, p. 97–103, 2013a. DOI: 10.3103/s1060992x13020082.

SUBBOTIN, S. Methods of sampling based on exhaustive and evolutionary search. **Automatic Control and Computer Sciences**, [*S. l.*], v. 47, n. 3, p. 113–121, 2013b. DOI: 10.3103/s0146411613030073.

SUYKENS, J. A.; VANDEWALLE, J. Least squares support vector machine classifiers. **Neural Processing Letters**, [*S. l.*], v. 9, n 3. p. 293–300, 1999. DOI: 10.1023/A:1018628609742.

TOMEK, I. An experiment with the edited nearest-neighbor rule. **IEEE Transactions on Systems, Man and Cybernetics**, [*S. l.*],v. 6, p. 448–452, 1976. DOI: 10.1109/TSMC.1976.4309523.

WILSON, D. L. Asymptotic properties of nearest neighbor rules using edited data. **IEEE Transactions on Systems, Man, Cybernetics**, [*S. l.*], v. 2, n. 3, p. 408–421, 1972. DOI: 10.1109/TSMC.1972.4309137.

WILSON, D. R.; Martinez, D. R. Reduction techniques for instancebased learning algorithms. **Machine Learning**, [*S. l.*],v. 38, n. 3, p. 257–286, 2000. DOI: 10.1023/A:1007626913721.
WINARNO, E.; HADIKURNIAWATI, W.; NIRWANTO, A.; ABDULLAH, D. Multi-view faces detection using viola-jones method. **Journal of Physics Conference Series**, [*S. l.*], v. 1114, n. 1, 2018. DOI: 10.1088/1742-6596/1114/1/012068.

YOON, B. J. Hidden Markov models and their application in the analysis of biological sequences. **Current Genomics**, [*S. l.*], v. 10, n. 6, p. 402–415, 2009. DOI: 10.2174/138920209789177575.

CRediT Author Statement

Reconhecimentos: Agradecemos à National University «Yuri Kondratyuk Poltava Polytechnic».

Financiamento: Não aplicável.

Conflitos de interesse: Não há conflitos de interesse.

Aprovação ética: Não aplicável.

Disponibilidade de dados e material: Não aplicável.

Contribuições dos autores: **Alla KAPITON**: Análise e interpretação dos dados. **Nataliia KONONETS**: Concepção, ideação, redação e revisão. **Volodymyr MOKLIAK**: Análise e interpretação dos dados. **Valentyna ONIPKO**: Coleta de dados. **Serhiy DUDKO**: Coleta de dados. **Vadym PYLYPENKO**: Colaboração na redação e revisão do artigo. **Anna SOKIL**: Colaboração na redação e revisão do artigo.

Processamento e editoração: Editora Ibero-Americana de Educação.
Revisão, formatação, normalização e tradução.