

# Estudo Comparativo de Abordagens Bioinspiradas para a Verificação de Impressões Digitais

Wesley Gomes de Brito Soares, Matheus Giovanni Pires  
 Universidade Estadual de Feira de Santana  
 Feira de Santana, Bahia, Brasil  
 e-mail: wesley.gbs@gmail.com, mgpires@ecom.uefs.br

**Resumo**—A utilização de senhas pessoais é amplamente utilizada em vários tipos de aplicações reais, tais como, caixas eletrônicas de bancos, no uso de cartões de débito ou crédito, dentre outros. Porém, o aumento do furto de tais senhas tem sido um grande problema para os usuários. Uma das formas para se tentar solucionar este problema é o uso de informações biométricas, como por exemplo, dados da impressão digital ou íris. Este artigo apresenta um estudo comparativo entre duas abordagens bioinspiradas aplicadas na resolução do problema da verificação de usuários, mais precisamente através do uso de impressões digitais. A primeira abordagem é baseada no comportamento forrageiro das formigas na busca por alimentos, conhecida como Colônia de Formigas, e a segunda é baseada nos mecanismos da seleção natural, conhecida como Algoritmos Genéticos. Os resultados obtidos são avaliados através de curvas ROC, permitindo analisar o desempenho das duas abordagens na resolução do problema da verificação de impressões digitais.

## I. INTRODUÇÃO

A utilização de senhas pessoais é a forma mais utilizada em sistemas que requerem algum nível de segurança para a identificação de usuários. Uma aplicação prática bem conhecida é a utilização de caixas eletrônicas dos bancos para efetuar saques ou consultas de saldo de contas correntes, por exemplo. Neste tipo de operação, o usuário deve informar além de sua senha pessoal, o número de sua agência e o número da conta corrente. Dependendo da operação a ser realizada, mais senhas poderão ser exigidas pelo sistema do banco. Há também o uso de cartões magnéticos que permitem a identificação de uma pessoa por meio da leitura dos mesmos. O uso de cartões magnéticos é comum em crachás de funcionários, onde o acesso a determinados lugares do seu local de trabalho só é permitido após a leitura do cartão do funcionário.

Um outro exemplo bem conhecido da utilização de senhas pessoais é o cartão de crédito. Para se utilizar um cartão de crédito o usuário necessita de uma senha pessoal. Com o aumento das compras pela Internet, o furto das senhas pessoais em sites de comércio eletrônico tornou-se muito comum. Para a Associação Brasileira das Empresas de Cartão de Crédito e Serviços (ABECS), a popularidade desse tipo de golpe também se apoia na impossibilidade de clonar os cartões de crédito com *chip*, que foi introduzido nos cartões de crédito e débito ao longo dos últimos anos. No 1º trimestre de 2016 os brasileiros movimentaram R\$ 269 bilhões com cartões de crédito e débito, crescimento de 7,2% em relação aos três primeiros meses de 2015. Os cartões de crédito registraram R\$ 165 bilhões (alta

de 3,8%), e os cartões de débito, R\$ 104 bilhões (alta de 13%) [1], tornando este tipo de golpe cada vez mais atraente.

Para contornar esses problemas, técnicas foram desenvolvidas para a identificação automática das pessoas por meio de suas características físicas ou comportamentais, ou seja, através da biometria [2]. Existem várias características biológicas que podem ser utilizadas no processo de identificação, as principais são: impressão digital, retina, íris, geometria da mão, face, voz e assinatura. O uso de biometria garante maior confiabilidade em relação ao uso de senhas, pois duas pessoas não possuem características idênticas e é quase impossível fraudar esse tipo de informação. No Brasil, várias empresas públicas e privadas têm utilizado a biometria para melhorar o controle e evitar fraudes, dentre elas destacam-se:

- DETRANS: projeto para interligar as escolas de formação de condutores aos servidores do DETRAN, onde os alunos registram sua presença pela biometria, armazenando as informações diretamente em um servidor externo, evitando as fraudes que ocorrem atualmente em relação ao cumprimento da carga-horária de aulas;
- Planos de saúde: as liberações de consultas são realizadas mediante a identificação do usuário no consultório médico, evitando o uso de cartão/senha por terceiros e, conseqüentemente, reduzindo os custos das operadoras;
- Bancos: os caixas eletrônicos utilizam a identificação de clientes por meio da impressão digital;
- Empresas privadas: para identificação e restrição de acessos, bem como registro do cartão de ponto dos colaboradores.

O mercado de biometria está em crescimento e a cada ano que passa a movimentação financeira aumenta em bilhões de dólares. Essa expansão deve-se, em grande parte, à diminuição do custo relacionado à implantação de um sistema biométrico, o qual está diminuindo com o decorrer do tempo. Somado a esse fator, o mercado exige/necessita cada vez mais de sistemas que forneçam segurança e confiabilidade, a fim de tornar as companhias menos vulneráveis a roubos de informações.

Considerando todos os fatores apresentados que evidenciam a grande utilização de dados biométricos para a o desenvolvimento de sistemas de segurança, este trabalho propõe um estudo comparativo entre duas abordagens, uma baseada em Colônia de Formigas e outra em Algoritmos Genéticos, para a resolução do problema da verificação de impressões digitais. A partir de uma base de dados de impressões digitais

previamente cadastrada, o sistema deverá verificar se existe uma impressão digital compatível com a impressão digital do usuário que se deseja identificar.

Este artigo está organizado da seguinte maneira: Na seção II são apresentados conceitos referentes à Biometria e Computação Bioinspirada, mais precisamente, sobre Colônia de Formigas e Algoritmos Genéticos. Em relação à Biometria, as vantagens e desvantagens de sua utilização são apresentadas, além de características biométricas que podem ser consideradas em sistemas de identificação, destacando-se a impressão digital. Além disso, os algoritmos baseados nas abordagens bioinspiradas são detalhados. Na seção III é descrita a utilização da abordagem baseada em Colônia de Formigas e Algoritmos Genéticos para a resolução do problema da verificação de impressões digitais. Na seção IV, são descritos os testes realizados e os resultados obtidos. Por fim, na seção V, as conclusões são apresentadas.

## II. FUNDAMENTAÇÃO TEÓRICA

Esta seção tem por objetivo apresentar os conceitos básicos de Biometria e sobre os algoritmos bioinspirados usados neste trabalho, Colônia de Formigas e Algoritmos Genéticos. Exemplos de aplicação destes algoritmos serão detalhados para um melhor entendimento dos mesmos no desenvolvimento do trabalho proposto neste artigo.

### A. Biometria

A Biometria se refere a identificação de pessoas baseadas em características físicas ou comportamentais, tais como, impressão digital, retina, íris, geometria da mão, face, voz, assinatura, dentre outros. A Biometria é mais confiável que os métodos tradicionais de identificação, pois essas técnicas são propensas a fraudes, como por exemplo, os cartões de crédito ou débito podem ser roubados e as senhas descobertas. Por outro lado, as características biológicas não podem ser esquecidas, compartilhadas ou extraviadas, requerendo ainda que a pessoa esteja presente para fornecer sua medida biométrica no instante da autenticação [2]. Para que uma característica física ou comportamental possa ser utilizada para a identificação de um indivíduo é necessário que ela atenda à quatro requisitos [3] [4]:

- **Universalidade:** toda pessoa deve possuir essa característica;
- **Unicidade:** dois indivíduos não devem possuir o mesmo padrão;
- **Permanência:** a característica deve permanecer inalterada ao longo do tempo;
- **Coletabilidade:** indica que é possível medir a característica de um modo quantitativo.

Para o desenvolvimento de um sistema biométrico, além das restrições supracitadas, a característica analisada deve atender a outros quesitos [4]:

- **Desempenho:** se refere à precisão e a velocidade alcançada no reconhecimento, os recursos necessários para atingí-los, bem como os fatores ambientais e operacionais que os afetam;

- **Aceitabilidade:** indica o grau em que as pessoas estão dispostas a aceitar um sistema biométrico em suas atividades diárias;
- **Evasão:** aponta a facilidade do sistema ser enganado por métodos fraudulentos.

Um sistema biométrico é um sistema de reconhecimento de padrões que opera sobre uma informação biométrica de um indivíduo, extraindo um conjunto de características desta informação e comparando-o com um segundo conjunto de características, o qual pode estar armazenado em um banco de dados. A depender da aplicação, um sistema pode ser definido como de *verificação* ou *identificação* [4]. Em um sistema de verificação, a entrada do sistema é uma consulta a uma informação biométrica associada com uma identidade (ID). Neste caso, o sistema verifica se a entrada é consistente ou não com a informação biométrica armazenada. Este tipo de sistema é ilustrado na Figura 1. Em um sistema de identificação, a entrada do sistema é somente uma consulta a uma informação biométrica (ver Figura 2). Neste caso, o sistema tenta responder a seguinte questão: Existe alguma informação biométrica na base de dados que se pareça com o dado consultado? A resposta é uma pequena lista contendo os modelos com maior grau de similaridade com a entrada.

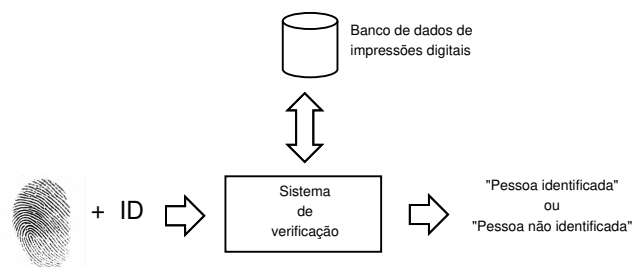


Figura 1: Sistema biométrico para verificação de impressões digitais.

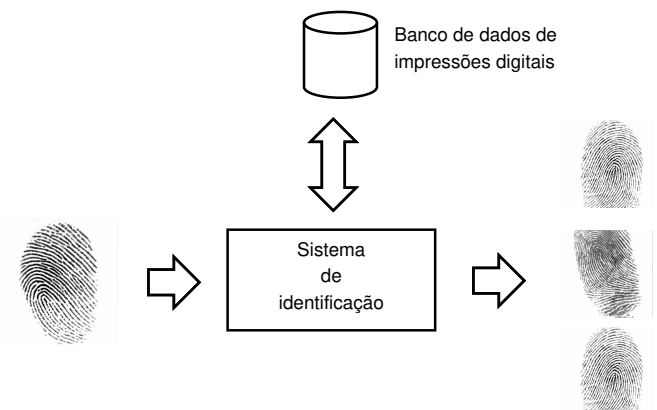


Figura 2: Sistema biométrico para identificação de impressões digitais.

A Biometria vem sendo empregada em diversas aplicações, podendo ser dividida em três grupos principais [4]:

- **Aplicações comerciais:** autenticação de usuários, segurança de dados, *e-commerce*, caixa eletrônico, cartões de

crédito, entre outras;

- **Aplicações governamentais:** carteira de identidade, licença de motorista, seguridade social, controle de passaporte, entre outras;
- **Aplicações forenses:** identificação de cadáveres, investigações criminais, identificação de terroristas, entre outras.

### B. Padrões Biométricos

Diversas características físicas e/ou comportamentais podem ser utilizadas em sistemas biométricos. Algumas já estão em uso enquanto outras ainda estão em fase de pesquisa por serem de alta complexidade. Nesta seção serão apresentadas algumas características biométricas que estão em uso e/ou pesquisa.

#### Análise de DNA

O DNA constitui um código único para cada indivíduo, exceto quando trata-se de gêmeos. As bases que o constituem podem se organizar em mais de 3 bilhões de formas [5]. Portanto, pode ser considerado um dos métodos mais precisos e difíceis de burlar. Atualmente, é utilizado somente em contexto forense, pois três quesitos limitam seu uso para outras aplicações [4]:

- 1) É fácil roubar um pedaço de DNA de um indivíduo inocente e usá-lo para outros fins;
- 2) A verificação de DNA requer o uso de métodos químicos que dependem de especialistas para realizá-lo, além de tratar-se de uma técnica invasiva;
- 3) Informações sobre susceptibilidade a doenças podem ser obtidas pela análise do DNA, o que pode promover o uso da informação genética para discriminação.

#### Geometria da face

A face humana é complexa e apresenta uma série de imperfeições, tornando-se um problema interessante a ser resolvido. Detalhes dos traços característicos de um indivíduo, pelos faciais e uso de acessórios (óculos escuros, por exemplo) fazem com que existam diversas pesquisas sobre o uso desta característica [6]. A biometria da face vem sendo empregada em aplicações de vigilância, com o uso de câmeras, para separar um indivíduo da multidão [7].

As abordagens usadas para reconhecimento facial geralmente baseiam-se em [4]:

- Localização e forma de atributos faciais, tais como, olhos, sobrancelhas, nariz, entre outros, e a relação espacial entre eles;
- Análise global de uma imagem que representa uma face como uma combinação ponderada de um número de faces canônicas.

#### Íris

A íris é a região anular do olho limitado pela pupila e a esclerótica (branco do olho). É formada durante o desenvolvimento do feto e estabilizada durante os dois primeiros anos de vida. A informação é única e muito distinta, ou seja, mesmo gêmeos possuem íris diferentes, assim como em impressões digitais. Além disso, é extremamente difícil falsificá-la, tanto por meio de cirurgias como através do uso de lentes de contato [4].

### Impressão Digital

A impressão digital de um indivíduo é única e permanece inalterada ao longo de sua vida. Ela é formada por cristas, as quais são definidas como segmentos de curva. Além das cristas, há também os vales, que são as regiões entre duas cristas [8]. A figura 3 ilustra um exemplo de impressão digital. As cristas são os segmentos em preto e os vales são as regiões em branco.



Figura 3: Características de uma impressão digital.

Computacionalmente, uma impressão digital é definida por um conjunto de curvas apresentadas nas cristas. Estas curvas são características denominadas minúcias e se apresentam das mais variadas formas: ponto-final, bifurcação, fragmento de linha, buraco, triângulo e gancho (Figura 4) [8], [9]. Entretanto, os sistemas de identificação de impressões digitais costumam utilizar pontos-finais e bifurcações, já que as demais tendem a serem introduzidas nas imagens devido a ruídos, criando falsas minúcias. Pontos-finais são aqueles onde a curva termina e as bifurcações são os pontos onde a curva se divide [10].

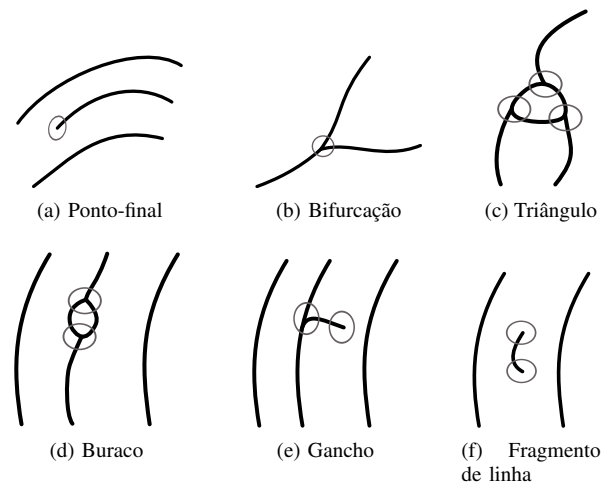


Figura 4: Exemplo de minúcias.

O reconhecimento de impressões digitais é um dos métodos mais usados dentre as técnicas de reconhecimento biométrico. Não é empregado apenas na aplicação das leis (identificação de suspeitos), mas também em aplicações comerciais que exigem um certo grau de segurança, como por exemplo o controle de acesso a locais restritos e transações financeiras.

### C. Otimização Contínua usando Colônia de Formigas

A otimização por Colônia de Formigas é uma meta-heurística para a solução de problemas combinatórios baseados no comportamento forrageiro das formigas na busca de alimentos [11]. Muitas espécies de formigas são quase cegas e a comunicação entre elas se dá através de uma substância química chamada feromônio. O uso dessa substância permite que as formigas encontrem o caminho mais curto entre uma fonte de alimento e o formigueiro. A figura 5 ilustra esse processo.

Entre o ninho e a fonte de alimentos há dois caminhos de tamanhos diferentes e não há nenhum feromônio depositado. Inicialmente, as formigas se espalharão em busca do alimento e, enquanto caminham, depositam no solo uma quantidade de feromônio. O caminho de menor tamanho permitirá que as formigas depositem mais feromônio, ao longo de idas e vindas entre o ninho e a fonte de alimentos. Como o seu deslocamento é baseado em trilhas de feromônio previamente depositados por outras formigas, elas serão motivadas a seguir pelo menor caminho.

Um algoritmo baseado na meta-heurística de Colônia de Formigas possui as seguintes características [12]:

- **Algoritmo não-determinístico:** inspirado no comportamento de formigas para determinação do caminho através de suas colônias para procura eficiente de uma fonte de alimento;
- **Algoritmo paralelo e adaptativo:** os agentes movem-se simultaneamente e de forma independente;
- **Algoritmo cooperativo:** cada agente constrói sua solução com base na informação deixada por outros agentes.

Um algoritmo baseado em Colônia de Formigas para a solução de problemas de otimização em domínios contínuos, denominado *ACO for continuous domains* ( $ACO_{\mathbb{R}}$ ) foi proposto por [13]. A ideia básica do  $ACO_{\mathbb{R}}$  é que a construção das soluções é feita de forma incremental pelas formigas, guiadas pelo feromônio e por uma Função de Densidade de Probabilidade (FDP).

No  $ACO_{\mathbb{R}}$  a informação do feromônio é armazenada em uma estrutura de dados chamada **arquivo de solução**, conforme ilustrado na figura 6. No arquivo de solução é armazenada uma quantidade  $k$  de soluções. Para cada solução  $s_l$  de um problema  $n$ -dimensional, o  $ACO_{\mathbb{R}}$  armazena neste arquivo os valores de suas  $n$  variáveis e o valor de sua função objetivo  $f(s_l)$ , a qual é responsável por medir a qualidade da solução. A  $i$ -ésima variável da  $l$ -ésima solução é denotada por  $s_l^i$ .

A execução do  $ACO_{\mathbb{R}}$  é descrito no algoritmo 1. O processo inicia-se com a definição dos parâmetros, como por exemplo, o número de formigas, e com a criação do arquivo de solução. Em seguida, inicia-se a etapa da construção das soluções, guiadas pelo feromônio e por uma FDP. A quantidade de soluções que serão criadas é igual ao número de formigas definidas. Por fim, os valores dos feromônios são atualizados. Este processo se repete até que um critério de parada seja atingido. A melhor solução para o problema será a que tiver a melhor função objetivo.

Maiores detalhes sobre o algoritmo  $ACO_{\mathbb{R}}$  podem ser encontrados em [13].

---

#### Algoritmo 1: Pseudocódigo do $ACO_{\mathbb{R}}$ [13].

---

```

1 Inicia os parâmetros;
2 enquanto critério de parada não atingido faça
3   | ControiSolução();
4   | AtualizaFeromônio();
5   | BuscaLocal() {opcional};

```

---

### D. Otimização usando Algoritmos Genéticos

Algoritmos Genéticos (AG) são métodos de otimização e busca baseados nos mecanismos da seleção natural introduzidos por [14]. Devido a grande capacidade de explorar espaços de busca grandes e irregulares, os AG têm sido bastante utilizados para estes tipos de problemas. Um Algoritmo Genético simples possui uma estrutura conforme o pseudocódigo ilustrado no algoritmo 2 [15].

---

#### Algoritmo 2: Pseudocódigo de um algoritmo genético simples.

---

```

1  $t \leftarrow 0$ ;
2 Inicializa  $P(t)$ ;
3 Avalia  $P(t)$ ;
4 enquanto critério de parada não atingido faça
5   |  $t \leftarrow t + 1$ ;
6   | Seleciona  $P(t)$  de  $P(t - 1)$ ;
7   | Altera  $P(t)$ ;
8   | Avalia  $P(t)$ ;

```

---

Durante a iteração  $t$ , o algoritmo genético mantém uma população de soluções candidatas (cromossomos). A primeira população, geralmente, é gerada de forma aleatória. Cada solução é avaliada para medir sua aptidão (ou *fitness*), ou seja, a qualidade da solução do problema representada por este cromossomo. Então, uma nova população (iteração  $t + 1$ ) é formada pela seleção favorável aos indivíduos mais aptos. Alguns membros desta nova população sofrerão alterações devido à ação dos operadores genéticos de cruzamento (*cross-over*) e mutação (*mutation*), enquanto outros permanecerão intactos. O cruzamento combina as características de dois cromossomos pais para formar dois cromossomos filhos. O objetivo da aplicação do cruzamento é trocar informações entre soluções em potencial, objetivando uma melhor exploração de uma determinada região do espaço de busca. Por outro lado, a mutação altera aleatoriamente um ou mais genes de um cromossomo selecionado, com o intuito de introduzir informação extra para a população. Em outras palavras, a mutação contribui com a diversidade da população, proporcionando a exploração de novos locais do espaço de busca e, conseqüentemente, evitando os máximos ou mínimos locais.

## III. METODOLOGIA

Nesta seção serão apresentadas as etapas realizadas para o desenvolvimento do projeto. Inicialmente será descrita uma visão geral do mesmo e, nos tópicos subsequentes, um maior detalhamento dos módulos envolvidos.

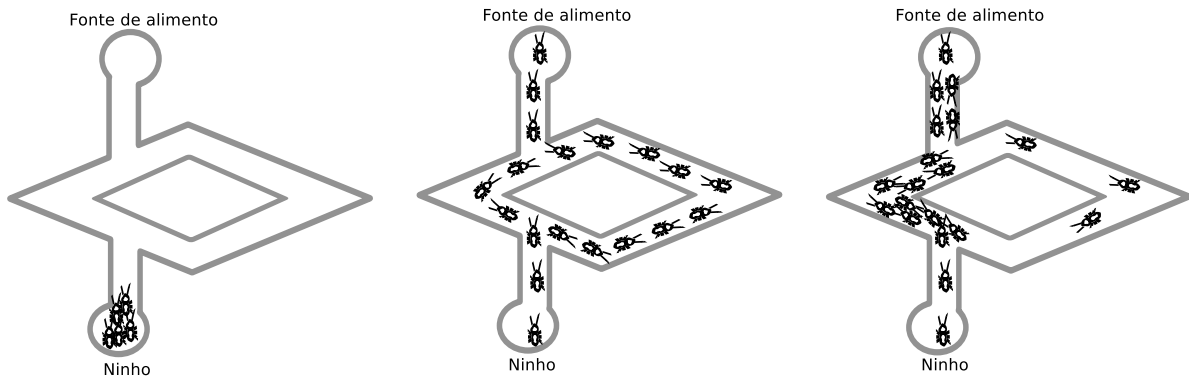


Figura 5: Construção da trilha de feromônio [11]

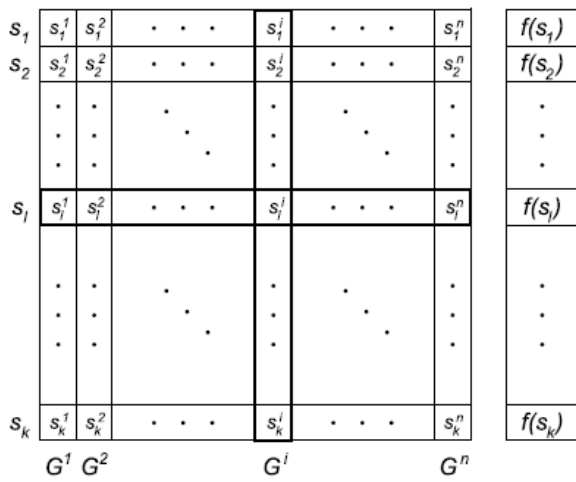


Figura 6: Representação do arquivo de solução [13].

#### A. Visão do projeto

O sistema para verificação de impressões digitais desenvolvido neste projeto é composto pelos seguintes módulos: pré-processamento das imagens, extração de minúcias, banco de dados e verificação das impressões digitais. A figura 7 mostra como estes módulos estão organizados.

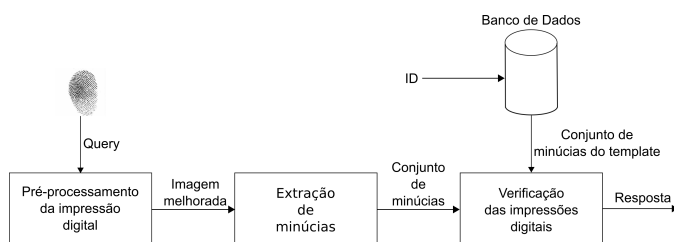


Figura 7: Organização de um sistema de verificação.

De acordo com a figura 7, a entrada do sistema é dada por duas informações: a impressão digital de um usuário (*query*) e um identificador (ID). O primeiro passo do processo de verificação é o pré-processamento da impressão digital do usuário, que tem por objetivo diminuir as imperfeições da

imagem e destacar as informações relevantes, neste caso, as cristas e vales. Em seguida, são extraídas as minúcias da impressão digital. De posse do conjunto de minúcias do usuário e do conjunto de minúcias do ID (armazenado no banco de dados), a comparação entre as impressões digitais é realizado por um algoritmo de verificação.

#### B. Pré-processamento das impressões digitais

Para assegurar que o desempenho de um sistema de verificação não seja comprometido devido a ruídos nas imagens, é necessário realizar um pré-processamento sobre as mesmas, com o objetivo de realçar as características que serão utilizadas pelo sistema.

A metodologia aplicada neste trabalho é a mesma apresentada por [8], na qual foram utilizadas as seguintes técnicas: segmentação, normalização, estimativa de orientação, estimativa da frequência das cristas, filtro de Gabor, binarização e, por fim, afinamento. A figura 8 ilustra o resultado da aplicação destas técnicas em uma impressão digital.

Neste trabalho foi utilizado os algoritmos desenvolvidos por Peter Kovesi, disponíveis em <http://www.peterkovesi.com/matlabfns/index.html#fingerprints>, para a realização da etapa de pré-processamento das impressões digitais.

#### C. Extração de minúcias

O módulo de extração de minúcias é responsável por detectar as minúcias das imagens. Neste trabalho foram extraídas apenas dois tipos de minúcias para a realização da verificação das impressões digitais, pontos finais e bifurcações.

A extração das minúcias é baseada no algoritmo desenvolvido por [8], chamado *Crossing Number (CN)*. Este algoritmo realiza uma varredura sobre a imagem no sentido anti-horário, verificando a vizinhança de cada *pixel* pertencente às cristas da impressão digital, por meio de uma matriz 3x3. Um exemplo desta matriz é ilustrada na figura 9. O valor CN para um dado *pixel P* é dado por:

$$CN = 0,5 \sum_{i=1}^8 |P_i - P_{i+1}|, \quad (1)$$

onde  $P_i$  é o valor do *pixel* na vizinhança de P.



Figura 8: Exemplo do pré-processamento de uma impressão digital [8].

Após este cálculo, um *pixel* será classificado como ponto-final, se CN for igual a um, ou bifurcação, se for igual a três.

Para cada minúcia extraída são armazenadas as seguintes informações: coordenadas  $x$  e  $y$ , ângulo de orientação da crista em que a minúcia está (obtido no passo anterior) e o tipo de minúcia (bifurcação ou ponto final).

A implementação desta etapa utilizou o código de Vahid K. Alilou, disponível em <http://www.mathworks.com/matlabcentral/fileexchange/44369-fingerprint-matching-a-simple-approach>.

P <sub>4</sub>	P <sub>3</sub>	P <sub>2</sub>
P <sub>5</sub>	P	P <sub>1</sub>
P <sub>6</sub>	P <sub>7</sub>	P <sub>8</sub>

Figura 9: Matriz de varredura da imagem [8].

#### D. Verificação de impressões digitais

O problema de verificação biométrica, em especial a impressão digital, pode ser formulado da seguinte forma. Sejam  $T$  e  $Q$  a representação do *template* e *query*, respectivamente, de forma que cada um deles é um vetor de minúcias:  $T = \{m_1^T, m_2^T, \dots, m_n^T\}$  e  $Q = \{m_1^Q, m_2^Q, \dots, m_p^Q\}$ , onde  $n$  e  $p$  é a quantidade de minúcias em  $T$  e  $Q$ , respectivamente. Cada minúcia é descrita pela 4-upla  $m = \{x, y, \theta, t\}$ , onde  $x$  e  $y$  são as coordenadas espaciais,  $\theta$  é o ângulo e  $t$  é o tipo da minúcia, podendo ser ponto-final ou bifurcação. O *template* representa o modelo de uma impressão digital de uma pessoa e a *query* é a entrada do sistema de verificação. A verificação de duas impressões digitais consiste da aplicação de duas etapas: (1) alinhamento e (2) comparação das minúcias. O alinhamento de duas impressões digitais é indispensável, já que permite maximizar o número de minúcias correspondentes na etapa seguinte. Este processo se dá através de reposicionamento das

minúcias da *query*, usando uma função de escala ( $s$ ), rotação e translação, dada pela equação 2 [10] [16]:

$$\text{map}(m^Q = \{x^Q, y^Q, \theta^Q, t^Q\}) = m^{Q'} = \{x^{Q'}, y^{Q'}, \theta^Q + \theta, t^Q\}, \text{ onde} \\ \begin{bmatrix} x^{Q'} \\ y^{Q'} \end{bmatrix} = s \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x^Q \\ y^Q \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \quad (2)$$

Em seguida, é realizada a comparação, onde duas minúcias  $m_i^T$  e  $m_j^{Q'}$  são consideradas correspondentes se a diferença espacial ( $\Delta e$ ) (equação 3) e diferença angular ( $\Delta \theta$ ) (equação 4) são inferiores a um limiar  $D_0$  e  $\theta_0$ , respectivamente [16]:

$$\Delta e(m_i^T, m_j^{Q'}) = \sqrt{(x_i^T - x_j^{Q'})^2 + (y_i^T - y_j^{Q'})^2} \leq D_0, \text{ e} \quad (3)$$

$$\Delta \theta(m_i^T, m_j^{Q'}) = \min(|\theta_i^T - \theta_j^{Q'}|, 360^\circ - |\theta_i^T - \theta_j^{Q'}|) \leq \theta_0 \quad (4)$$

Este cálculo é feito minúcia a minúcia, de forma que uma minúcia  $m_j^{Q'}$  pode ter apenas uma minúcia  $m_i^T$  correspondente. Após o processamento, obtém-se um conjunto  $C$  de minúcias correspondentes, com o qual é calculado o grau de similaridade, ou *score*, entre  $T$  e  $Q$ , conforme a equação 5 [16]. Se o *score* for maior que um limiar  $L$ , as impressões digitais pertencem ao mesmo indivíduo.

$$\text{score} = \frac{|C|}{\frac{n+p}{2}} \quad (5)$$

O problema da verificação entre duas impressões digitais consiste em alinhar as duas impressões, a fim de possibilitar a comparação das mesmas. Para que este alinhamento contribua com resultados satisfatórios, é necessário encontrar bons valores para as variáveis da equação 2. Este processo consiste em um problema de otimização, que neste trabalho será resolvido por duas abordagens bioinspiradas, Colônia de Formigas e Algoritmos Genéticos.

#### E. Modelagem do problema usando Colônia de Formigas

O arquivo de solução tem seus componentes dispostos conforme ilustrado na figura 10. Para inicializar esta estrutura são gerados valores iniciais aleatoriamente e igualmente espaçados, impedindo que duas ou mais soluções estejam associadas a uma mesma região. A função objetivo adotada é o grau de similaridade entre o *template* e a *query*, portanto, utilizou-se a equação 5.

$s$	$\theta$	$T_x$	$T_y$
-----	----------	-------	-------

Figura 10: Definição do arquivo de solução para verificação de impressão digital

#### F. Modelagem do problema usando Algoritmos Genéticos

A modelagem do Algoritmo Genético é bastante similar à do ACO<sub>R</sub>. A codificação cromossômica baseia-se na figura 10, com uso da representação real das variáveis, assim a manipulação da informação torna-se mais fácil. A população inicial é grande, e a cada geração dois novos filhos são obtidos através de um operador de cruzamento aritmético para

representação real. O operador de mutação é uma adaptação do *bit-flip* para a representação real, onde é aleatoriamente adicionado ou subtraído um valor ao gene escolhido. A função de *fitness* é exatamente igual a função objetivo definida pela equação 5.

#### IV. EXPERIMENTOS E RESULTADOS

Nesta seção serão apresentados os experimentos realizados para avaliar o desempenho das abordagens baseadas em Colônia de Formigas e Algoritmos Genéticos aplicadas ao problema da verificação de impressões digitais. Inicialmente, a base de imagens utilizada nos experimentos será apresentada, e em seguida, os parâmetros dos algoritmos são definidos. Por fim, os resultados obtidos são apresentados.

##### A. Descrição do experimento

Todos os experimentos foram executados sobre a base de dados DB1, subconjunto B, da *Fingerprint Verification Competition* (FVC), que ocorreu em 2002 [17]. Esta base de dados caracteriza-se por ter imagens adquiridas pelo sensor óptico *TouchView II*, cada uma delas com tamanho de 388x374 (142 Kpixels) e resolução de 500 dpi. Ela conta com dois conjuntos de impressões digitais denominados **A** e **B**. O primeiro possui 110 impressões digitais diferentes, cada uma com 8 amostras. O segundo é um subconjunto do primeiro, contendo 10 impressões digitais diferentes e 8 amostras de cada.

A forma de avaliação adotada neste trabalho é o mesmo utilizado na FVC, a qual é dividida em verificação de genuínos e de impostores. Na verificação de genuínos, cada amostra do conjunto é comparada às amostras remanescentes do mesmo dedo. Já na verificação de impostores, a primeira amostra de cada dedo é comparada à primeira amostra dos outros dedos. Dessa forma, haverá 280 comparações para os testes de genuínos e 45 para os impostores. Os resultados obtidos foram apresentados por curvas ROC [18]. As curvas ROC geradas expressam as taxas de verdadeiro positivo por falso positivo, possibilitando a análise do desempenho de cada abordagem usada neste trabalho.

Dado um conjunto de testes **T** em que cada impressão digital pertence a uma classe positiva ou negativa, ao processar cada impressão, um classificador produzirá uma saída positiva ou negativa, para diferentes limiares aplicados. Logo, há quatro possibilidades de classificação para cada impressão digital. Se uma impressão é positiva e o sistema a avalia como positiva, então é contabilizado um verdadeiro positivo; caso contrário, falso negativo. Se uma impressão é negativa e é avaliada como negativa, então é contabilizado um verdadeiro negativo; caso contrário, falso positivo [18]. As taxas de verdadeiro positivo (*vp*) e falso positivo (*fp*) são apresentadas na equação 6 e 7.

$$vp = \frac{\text{número positivos corretamente classificados}}{\text{total de positivos}} \quad (6)$$

$$fp = \frac{\text{número negativos incorretamente classificados}}{\text{total de negativos}} \quad (7)$$

Além das taxas de verdadeiro positivo e falso positivo utilizadas para a construção das curvas ROC, foi utilizada outra medida para avaliar o desempenho dos algoritmos, chamada AUC (*Area Under an ROC Curve*). A AUC é usada como medida de precisão de uma resposta do sistema e segue a interpretação apresentada na tabela I [19].

Tabela I: Interpretação da Área Sob a Curva ROC

AUC	Precisão do sistema
= 5	Sem precisão
$0,7 \leq AUC < 0,8$	Aceitável
$0,8 \leq AUC < 0,9$	Excelente
$\geq 0,9$	Ótima

##### B. Avaliação do $ACOR_{\mathbb{R}}$ e do AG

Os parâmetros do algoritmo de colônia de formigas estão definidos na tabela II. O valor de  $N_F$ ,  $q$  e  $k$  foram definidos a partir dos experimentos realizados em [13].

Tabela II: Parâmetros usados no  $ACO_{\mathbb{R}}$

Parâmetros	Símbolo	Valor
Tamanho do arquivo de solução	$k$	100
Localidade do processo de busca	$q$	0,001
Taxa de convergência	$\xi$	0,85
Número de formigas	$N_F$	15
Limiar de diferença angular	$\theta_0$	20

Os testes ocorreram assumindo os seguintes valores para  $D_0$ : 15, 20, 25 e 30. A figura 11 mostra a curva ROC para cada um desses valores. Nela é possível perceber que o  $ACO_{\mathbb{R}}$  tem desempenho semelhante em todos os casos, isso torna-se claro na tabela III, em que destaca-se um aumento máximo de 4% no desempenho do sistema, evidenciado pela AUC. Logo, para todos os valores testados, pode-se considerar que o sistema tem um desempenho aceitável. Porém, levando-se em consideração que aumentar  $D_0$  fará com o sistema seja mais flexível, ou seja, suscetível a erros, o melhor valor para esse parâmetro é 20.

Tabela III: Desempenho do  $ACO_{\mathbb{R}}$

$D_0$	Área sob a curva (AUC)
15	0,765595
20	0,801429
25	0,770913
30	0,802460

Os parâmetros do AG estão apresentados na tabela IV, os quais foram baseados em [20]. Os valores aplicados a  $D_0$  são os mesmos usados no  $ACO_{\mathbb{R}}$ . A figura 12 ilustra as curvas ROC obtidas pelos resultados do algoritmo genético. Analisando-as, percebe-se que, assim como no  $ACO_{\mathbb{R}}$ , todas as configurações obtiveram resultados semelhantes. Os resultados de AUC descritos na tabela V confirmam tal observação. Dos resultados de AUC obtidos, o melhor resultado foi alcançado quando  $D_0 = 20$ , ou seja, AUC de 0,799881.

Tabela V: Desempenho do AG

$D_0$	Área sob a curva (AUC)
15	0,755833
20	0,799881
25	0,763492
30	0,788452

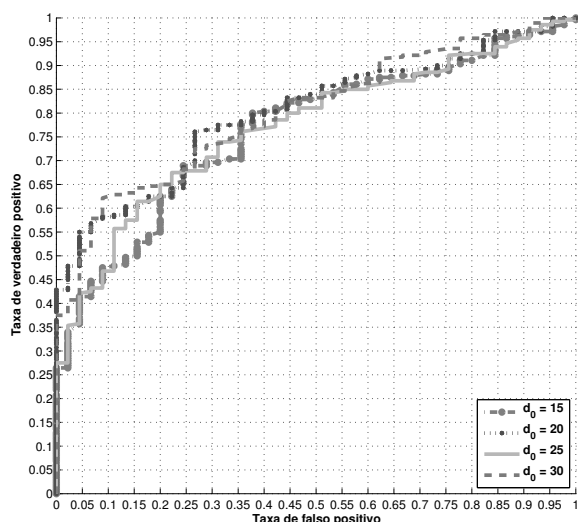
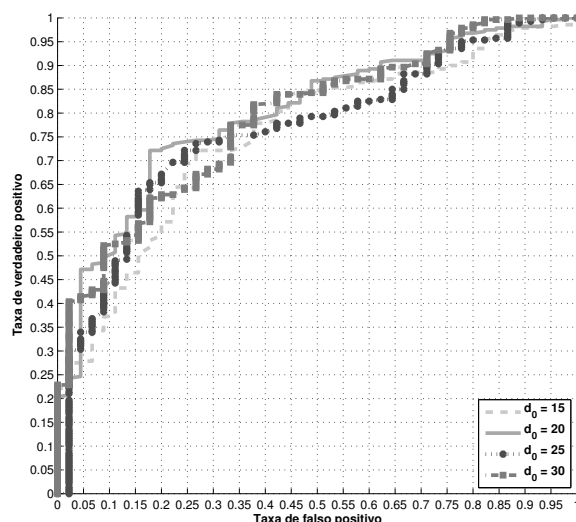
Figura 11: Curva ROC do ACO<sub>R</sub>

Figura 12: Curva ROC do AG

Tabela IV: Parâmetros usados no AG

Parâmetros	Símbolo	Valor
Número de gerações	$N_g$	100
Tamanho da população	$T_p$	100
Taxa de cruzamento	$C_x$	0.8
Taxa de mutação	$M_x$	0.2
Limiar de diferença angular	$\theta_0$	20

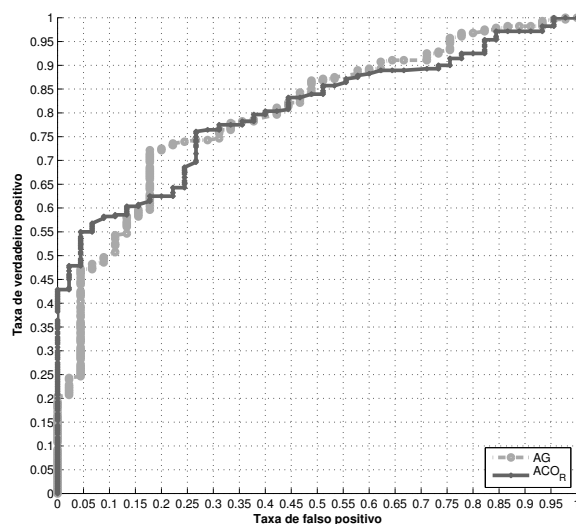
Ao comparar os melhores resultados obtidos pelos dois algoritmos observa-se que o ACO<sub>R</sub> tem um desempenho ligeiramente melhor que o AG. A figura 13 ilustra como as curvas ROC têm comportamento parecido.

## V. CONCLUSÃO

Neste artigo foi proposto um estudo comparativo entre duas abordagens bioinspiradas para a resolução do problema da verificação de impressões digitais. Mais precisamente, um Algoritmo Genético e um algoritmo de Colônia de Formigas foram usados para otimizar as variáveis de uma função, que é responsável em realizar o alinhamento das minúcias entre duas impressões digitais, para que as mesmas pudessem ser comparadas.

Os experimentos utilizaram a base de dados FVC 2002, a qual contém imagens com alto grau de distorções, tais como, rotação, translação, escala e ruídos. O processo consistiu na verificação das impressões genuínas, ou seja, imagens de um mesmo dedo são comparadas entre si, e na verificação das impressões impostoras, isto é, comparação entre imagens de dedos diferentes. Os resultados obtidos de cada algoritmo foram expressos por curvas ROC, possibilitando a análise do desempenho de cada abordagem usada neste trabalho.

Observando o critério da AUC as duas abordagens apresentaram resultados satisfatórios, com uma pequena vantagem para o ACO<sub>R</sub>. A fim de melhorar o desempenho, pode-se optar por implementar estratégias de otimização complementares às abordagens usadas, como o elitismo, também poderia-se aumentar o tamanho da população de cromossomos ou

Figura 13: Curva ROC do ACO<sub>R</sub> e AG com  $D_0 = 20$ 

do arquivo de solução, ou modificar qualquer um de seus atributos. Porém, maior impacto seria causado através da melhoria do processo de extração de minúcias, ou até mesmo o uso de outras características da impressão digital no processo de verificação.

## REFERÊNCIAS

- [1] Associação Brasileira das Empresas de Cartão de Crédito e Serviços. (2016) Cartões somam R\$ 269 milhões em compras no IT16, aponta Abecs. [Online]. Available: <http://www.abecs.org.br/noticia/cartoes-somam-r-269-bilhoes-em-compras-no-1t16-aponta-abecs>
- [2] A. K. Jain, S. Prabhakar, and S. Pankanti, "On the similarity of identical twin fingerprints," *Pattern Recognition*, vol. 35, pp. 2653–2663, 2002.
- [3] E. de Araújo Jardini, "Mfis: Algoritmo de reconhecimento e indexação em base de dados de impressões digitais em espaço métrico," Ph.D. dissertation, Universidade de São Paulo, Escola de Engenharia de São Carlos, Departamento de Engenharia Elétrica, 2007.



- [4] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, 2004.
- [5] Galileu, "Você é a senha," *Revista Galileu*, jul. 2004. [Online]. Available: [http://revistagalileu.globo.com/EditoraGlobo/componentes/article/edg\\_article\\_print/1,3916,754754-1938-1,00.html](http://revistagalileu.globo.com/EditoraGlobo/componentes/article/edg_article_print/1,3916,754754-1938-1,00.html)
- [6] W. J. B. Alves, "Identificação de pessoas através de algoritmo genético aplicado em medidas das proporções Áureas da face humana," Master's thesis, Escola de Engenharia de São Carlos da Universidade de São Paulo, 2008.
- [7] D. Fortes, "Biometria é a senha," *INFO online*, vol. 223, out. 2004a. [Online]. Available: [http://info.abril.com.br/edicoes/223/arquivos/4689\\_1.shl](http://info.abril.com.br/edicoes/223/arquivos/4689_1.shl)
- [8] R. Thai, "Fingerprint image enchacement and minutae extraction," *Honours Programme of the School of Computer Science and Software Engineering, The University of Western Australia*, 2003.
- [9] A. Jain, L. Hong, and R. Bolle, "On-line fingerprint verification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 302–314, 1997.
- [10] X. Tan and B. Bhanu, "Fingerprint matching by genetic algorithms," *Pattern Recognition*, vol. 39, pp. 465–477, 2006.
- [11] M. Dorigo, E. Bonabeau, and G. Theraulaz, *Swarm intelligence: from natural to artificial systems*. New York, NY, USA: Oxford University Press, Inc., 1999. [Online]. Available: <http://portal.acm.org/citation.cfm?id=328320>
- [12] L. dos Santos Coelho and R. F. T. Neto, "Colônia de formigas: Uma abordagem promissora para aplicações de atribuição quadrática e projeto de layout," in *XXIV ENEGEP Florianópolis, SC, Brasil*, 2004.
- [13] K. Socha and M. Dorigo, "Ant colony optimization for continuous domains," *European Journal of Operational Research*, vol. 185, no. 3, 2008.
- [14] J. H. Holland, *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press, 1975.
- [15] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs (3rd Ed.)*. London, UK, UK: Springer-Verlag, 1996.
- [16] A. K. Jain and D. Maltoni, *Handbook of Fingerprint Recognition*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2009.
- [17] BioLab - University of Bologna. (2002) Second international competition for fingerprint verification algorithms. [Online]. Available: <http://bias.csr.unibo.it/fvc2002/>
- [18] T. Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, jun 2006.
- [19] D. W. Hosmer and S. Lemeshow, *Applied logistic regression*, ser. Wiley series in probability and statistics. New York, Chichester, Weinheim: John Wiley & Sons, Inc. A Wiley-Interscience Publication, 2000.
- [20] M. G. Pires, F. V. Duarte, and A. Gonzaga, "Verificação de impressões digitais usando algoritmos genéticos," *II Workshop de Visão Computacional (WVC), São Carlos*, 2006.