inpi instituto nacional da propriedade industrial

(11) Número de Publicação: **PT 117136 B**

(51) Classificação Internacional:
G06F 21/16 (2013.01)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2021.03.23**

(30) Prioridade(s):

(43) Data de publicação do pedido: **2022.09.23**

(45) Data e BPI da concessão: **2024.09.20**
186/2024

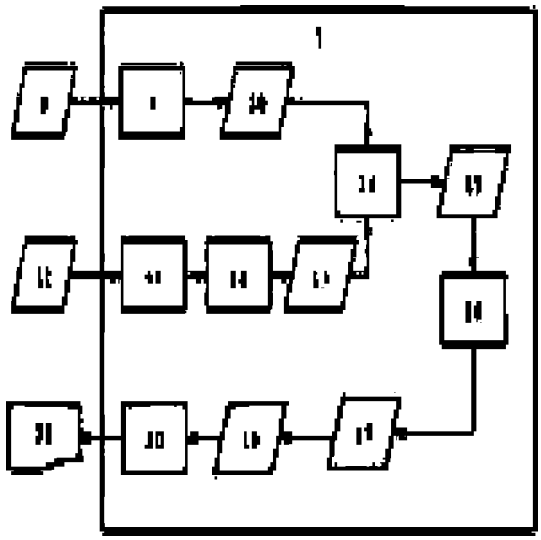
(73) Titular(es):
UNIVERSIDADE DE COIMBRA
PAÇO DAS ESCOLAS 3004-531 COIMBRA **PT**
INCM - IMPRENSA NACIONAL - CASA DA MOEDA, S.A. **PT**

(72) Inventor(es):
NUNO MIGUEL MENDONÇA DA SILVA GONÇALVES **PT**
FARHAD SHADMAND **IR**

(74) Mandatário:
TIAGO ANDRÉ DELGADO REIS NOBRE
ALAMEDA DOS OCEANOS, 41K-21 1990-207 PARQUE DAS NAÇÕES **PT**

(54) Epígrafe: **SISTEMAS DE CODIFICAÇÃO, DESCODIFICAÇÃO E VALIDAÇÃO DA INTEGRIDADE DE UM DOCUMENTO DE SEGURANÇA COM UMA IMAGEM CODIFICADA POR ESTEGANOGRAFIA E MÉTODOS, DOCUMENTO DE SEGURANÇA, APARELHOS COMPUTACIONAIS, PROGRAMAS DE COMPUTADOR E MEIOS DE LEITURA ASSOCIADOS**

(57) Resumo:



SISTEMA DE CODIFICAÇÃO DE UM DOCUMENTO DE SEGURANÇA COM UMA IMAGEM CODIFICADA POR ESTEGANOGRAFIA RESISTENTE À IMPRESSÃO (1) COMPREENDENDO UM MÓDULO CODIFICADOR PARA CORREÇÃO DE ERROS BINÁRIO E INTRODUÇÃO DE REDUNDÂNCIA (9), QUE CONVERTE MENSAGENS SECRETA (8) EM MENSAGENS BINÁRIA A CODIFICAR (10); UM MÓDULO DE DETEÇÃO FACIAL, ALINHAMENTO, RECORTE E REDIMENSIONAMENTO (12), QUE PROCESSA IMAGENS PARCIAIS A CODIFICAR (13) A PARTIR DE IMAGENS INTEGRAIS A CODIFICAR (11); UM MÓDULO CODIFICADOR (14), QUE ORIGINA IMAGENS PARCIAIS CODIFICADAS (15); UM MÓDULO INTEGRADOR DE IMAGENS (16), QUE ORIGINA IMAGENS INTEGRAIS CODIFICADAS (17); UM MÓDULO (20) DE TRANSMISSÃO FÍSICA PARA DOCUMENTOS DE SEGURANÇA (21); UM MÓDULO DE PRÉ-PROCESSAMENTO DE IMAGENS (48); E UM SISTEMA DE DESCODIFICAÇÃO RELACIONADO, OPERANDO POR REDES NEURONAIS, TREINADAS COM BASE EM IMAGENS DE TREINO COM RUÍDO SIMULADO (19). A INVENÇÃO PERMITE ESCONDER MENSAGENS SECRETAS EM IMAGENS FACIAIS DE DOCUMENTOS DE SEGURANÇA, NOMEADAMENTE DOCUMENTOS DE IDENTIFICAÇÃO CIVIL E DE VIAGEM.

RESUMO

"SISTEMAS DE CODIFICAÇÃO, DESCODIFICAÇÃO E VALIDAÇÃO DA INTEGRIDADE DE UM DOCUMENTO DE SEGURANÇA COM UMA IMAGEM CODIFICADA POR ESTEGANOGRAFIA E MÉTODOS, DOCUMENTO DE SEGURANÇA, APARELHOS COMPUTACIONAIS, PROGRAMAS DE COMPUTADOR E MEIOS DE LEITURA ASSOCIADOS"

Sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) compreendendo um módulo codificador para correção de erros binário e introdução de redundância (9), que converte mensagens secreta (8) em mensagens binária a codificar (10); um módulo de deteção facial, alinhamento, recorte e redimensionamento (12), que processa imagens parciais a codificar (13) a partir de imagens integrais a codificar (11); um módulo codificador (14), que origina imagens parciais codificadas (15); um módulo integrador de imagens (16), que origina imagens integrais codificadas (17); um módulo (20) de transmissão física para documentos de segurança (21); um módulo de pré-processamento de imagens (48); e um sistema de descodificação relacionado, operando por redes neuronais, treinadas com base em imagens de treino com ruído simulado (19).

A invenção permite esconder mensagens secretas em imagens faciais de documentos de segurança, nomeadamente documentos de identificação civil e de viagem.

DESCRIÇÃO

"SISTEMAS DE CODIFICAÇÃO, DESCODIFICAÇÃO E VALIDAÇÃO DA INTEGRIDADE DE UM DOCUMENTO DE SEGURANÇA COM UMA IMAGEM CODIFICADA POR ESTEGANOGRAFIA E MÉTODOS, DOCUMENTO DE SEGURANÇA, APARELHOS COMPUTACIONAIS, PROGRAMAS DE COMPUTADOR E MEIOS DE LEITURA ASSOCIADOS"

DOMÍNIO TÉCNICO

A presente invenção insere-se na área de esteganografia, em particular nos sistemas e métodos de esteganografia resistentes à impressão e à captura de imagem por uma câmara, na língua inglesa designada por *printer-proof steganography*.

TÉCNICA ANTERIOR

A integridade das imagens de faces em documentos de segurança, por exemplo documentos de identificação civil e de viagem, nomeadamente os cartões de identidade nacional e os passaportes, é frequentemente condicionada pela falsificação dos respetivos documentos. Os documentos de segurança, que estão incluídos no âmbito da presente invenção, podem ser identificados como ID-MRTD (*Identification Documents and Machine-Readable Travel Documents*).

Dadas as suas características, a imagem da face de um indivíduo é globalmente reconhecida como uma forma primordial de identificação de pessoas, constituindo-se como um dos principais elementos biométricos, cada vez mais comuns

e cada vez mais usados para o propósito mencionado.

A imagem da face é, pois, também considerada como um dos elementos de segurança mais falsificados e atacados nos documentos de identificação civil e de viagem. Uma das formas existentes de garantir a integridade das imagens das faces destes documentos é através da alteração da imagem por forma a codificar nela uma informação escondida, com o objetivo de usar esta informação por sistemas de verificação da sua integridade, quer manuais quer automáticos, cruzando a informação escondida na imagem e descodificada ou obtida no momento da verificação com a informação esperada para o contexto em que a imagem se insere. Um exemplo deste mecanismo de segurança é codificar na imagem da face de uma pessoa constante no seu passaporte o número do passaporte relacionado. Será, pois, fácil a um sistema de verificação cruzar o conteúdo escondido na imagem com o conteúdo do contexto dessa imagem, assim formulando um juízo relativo à integridade da foto e do documento.

Em contextos afins, a necessidade de verificação dos documentos estende-se a outros elementos para além da face da pessoa, por exemplo, os elementos do design impressos no cartão e que sejam comuns a todos os cartões. É então comum confirmar a integridade do próprio documento de suporte através da análise das imagens nele constantes.

No contexto da presente invenção, nomeadamente na área de esconder uma informação codificada numa imagem, a esteganografia é uma técnica para esconder a existência de informação privada numa informação de suporte. Desta forma, a simples análise ou a visualização da informação de suporte não é suficiente nem para revelar a existência de informação escondida, nem para a descodificar. Do ponto de vista da tipologia, a informação a esconder pode ser alfanumérica,

binária ou matricial (imagem) e a informação de suporte pode ser igualmente alfanumérica, binária ou matricial (imagem).

Portanto, a esteganografia é uma técnica muito usual para esconder informação em imagens, de forma a disfarçar a própria existência de informação escondida. A ideia principal da esteganografia é a integração da informação a esconder na informação visual de base, obrigando necessariamente a uma perda de informação desta última, para poder acomodar a informação da primeira. O sucesso de um método de esteganografia reside na capacidade de esse método manter o máximo possível das propriedades visuais da imagem de base, com uma perda limitada e tão reduzida quanto possível da perceção daquela.

PROBLEMAS DA TÉCNICA ANTERIOR

O método mais primitivo e simples de esteganografia passa por substituir o bit menos significativo dos pixéis de uma imagem por um bit da informação a esconder. A informação escondida é, neste caso, a concatenação de todos os bits menos significativos da imagem final. Este método simples tem um impacto muito reduzido na perceção da imagem de base na imagem final, pois a informação perdida, ou destruída, é muito pouca, correspondendo em termos gerais a uma pequena redução da resolução da imagem real e uma alteração de algumas componentes de alta frequência, pouco impactantes na perceção global da imagem. Apesar desta simplicidade na codificação e descodificação da imagem, este método não apresenta elevada segurança na proteção dos dados escondidos e, sobretudo, não é resistente a ruído nem a algumas manipulações habituais da imagem, tais como compressão (por exemplo no formato JPEG, em que a sigla vem do nome em língua

inglesa "*Joint Photographic Experts Group*", que é o grupo responsável pelo padrão JPEG.

Existem dezenas de métodos de esteganografia mais sofisticados que permitem esconder a informação e protegê-la de forma mais segura.

Contudo, os métodos conhecidos de esteganografia não são resistentes à transmissão da informação por canal físico, mantendo as suas propriedades apenas em canais digitais, geralmente sem perdas. Efetivamente, quando há lugar a uma mudança do formato digital para o formato físico, sendo tal processo designado por transdução, e posteriormente uma reversão para o formato digital, existe, no caso geral, uma perda significativa de informação que não preserva a informação escondida pelo método de esteganografia.

Os métodos de transdução digital-física mais comuns são os executados por dispositivos de transdução, como por exemplo os scanners, as câmaras, as retinas artificiais, as impressoras e os ecrãs. Frequentemente, a transdução de imagens ocorre por meio do processamento dos pixéis, considerando que uma imagem possui uma representação ótica, que pode ser sentida diretamente ou indiretamente por sinais óticos e uma representação de dados baseados em sinais eletrónicos, tais como a voltagem, a corrente elétrica ou a carga elétrica, tanto em forma analógica, quanto em forma digital. Portanto, um pixel pode ser representado oticamente ou eletronicamente.

No caso da impressão da imagem, esta impressão pode ser realizada por qualquer tipo de impressão em qualquer suporte físico, nomeadamente o papel; os polímeros, como por exemplo o policarbonato e o policloreto de vinila; as chapas metálicas; as resinas; os vernizes; as fitas; os materiais holográficos ou qualquer outro suporte físico. No caso da

visualização por ecrã, também podem ser usados diversos tipos de equipamentos eletrónicos, nomeadamente os de tubo de raios catódicos (CRT), os ecrãs de cristal líquido (LCD), os cristais líquidos, os ecrãs de díodos emissores de luz (LED) e todas as restantes tecnologias afins ou derivadas. Por outro lado, no processo de transdução físico-digital, a imagem é captada por dispositivos óticos, nomeadamente câmara fotográficas ou de vídeo, usando qualquer tipo de tecnologia de visão.

Assim, a imagem digital, que foi recuperada de uma imagem impressa em suporte físico e que tenha sido codificada por um processo de esteganografia comum, não preserva a informação escondida, tendo esta sido destruída totalmente ou em grande parte.

Adicionalmente, importa referir que apesar de a captação da imagem impressa ou visualizada num ecrã poder ser efetuada por um qualquer dispositivo de visão, é de particular interesse que esta captação possa ser efetuada por dispositivos móveis ubíquos, em especial os telemóveis pessoais, em inglês *smartphones*, uma vez que a verificação da integridade dos documentos ID-MRTD pode, neste caso, ser efetuada por agentes da autoridade ou pelos próprios cidadãos em qualquer momento, e não só em portais ou dispositivos fixos.

A circunstância da solução ser utilizável com efetividade por dispositivos móveis ubíquos coloca dois desafios adicionais ao problema técnico a resolver. Por um lado, os mecanismos matemáticos e formais de solução do problema terão que se basear em algoritmos cujo custo computacional seja comportável nos dispositivos móveis, assim produzindo um juízo sobre a integridade do documento em tempo aceitável, preferencialmente abaixo de 1 segundo, mas aceitável até poucos segundos. Por outro lado, para que

a aplicação prática de uma tecnologia que se proponha a solucionar os problemas do estado da técnica seja efetiva, é necessário que a validação da integridade do documento possa ser efetuada num modo desconectado de uma rede de comunicação, isto é, num modo *offline* e sem ter que necessariamente recorrer a sistemas centrais remotos, evitando que o serviço possa ficar indisponível por falta de conectividade ou por existência de conectividade comprometida do ponto de vista da segurança.

De qualquer modo, a contextualização dos problemas do estado da técnica não implica em que não existam situações em que seja preferível o modo de conexão com uma rede de comunicação, isto é, o modo *online*, em que o dispositivo móvel comunica com o sistema central remoto.

Resumindo, é conveniente que a tecnologia permita uma validação da integridade de documentos de segurança em modos *offline* e *online*, cuja arquitetura efetiva seja desenhada em função do caso específico e concreto para o qual se destina, sendo ainda conveniente que o sistema, qualquer que seja o modo adotado, seja o *offline* ou o *online*, cumpra todos os requisitos de segurança aplicacional que impeça o sistema de ser violado.

Deste modo, o que a presente invenção pretende resolver é o problema de codificar uma imagem usando esteganografia resistente à alteração de meio físico ou transmissão física com perdas. Estes métodos podem ser, por conveniência linguística, designados por esteganografia resistente à impressão ou, na língua inglesa, *printer-proof steganography method*.

SOLUÇÃO DO PROBLEMA

A presente invenção visa solucionar os problemas

do estado da técnica por intermédio da utilização da tecnologia de redes generativas adversárias para gerar imagens codificadas incorporáveis em documentos de segurança com informação secreta escondida, com alterações mínimas em relação às imagens originais, e que mantenham, quer a capacidade do sistema de descodificar a informação escondida após a transmissão para um meio físico, quer a capacidade de perceção por seres humanos ou sistemas automáticos de reconhecimento das imagens presentes nos documentos de segurança.

As redes generativas adversárias, conhecidas na língua inglesa por "*generative adversarial networks*", também conhecidas por GANs, são redes neuronais cuja arquitetura assenta em duas subredes adversárias, tal como na teoria dos jogos: a rede geradora e a rede discriminadora. O objetivo da rede geradora é criar exemplos de imagens, nomeadamente imagens artificiais, que se assemelhem a um tipo ou estilo de imagens dada como iniciais. Contrariamente, a rede discriminadora tem como objetivo distinguir as imagens reais das imagens geradas artificialmente. Assim, para que a rede generativa adversária tenha a capacidade de gerar imagens artificiais com elevado realismo é, pois, tal como em qualquer sistema de aprendizagem por máquina, necessário treinar a rede com um número elevado de exemplos. Tal arquitetura é então treinada a partir de um conjunto de imagens de entrada, um banco de dados, *dataset* em inglês, de tal forma que os pesos e os parâmetros da rede são ajustados até que a rede geradora seja capaz de gerar imagens artificiais que a rede discriminadora não seja capaz de distinguir. No fim da fase de treino, a rede geradora adversária é então capaz de gerar imagens artificiais com elevado realismo e que são facilmente percebidas como imagens reais por seres humanos.

No que se refere à geração das imagens codificadas, a rede geradora engloba em si as capacidades de codificação, mas também as de descodificação, a fim de testar a capacidade da rede geradora na criação de imagens com mensagens secretas e cuja informação possa ser posteriormente descodificada com sucesso.

Logo, um contributo técnico relevante proporcionado pela presente invenção está relacionado com o facto de que o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) e o sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) sejam treinados para manter a resistência da informação codificada à descodificação, na presença de um número muito elevado de fatores físicos e digitais de distorção, erro e perda de informação. Para tal, é incorporado um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3), o qual inclui uma rede neuronal do tipo generativa adversária geradora (6) e uma rede neuronal do tipo generativa adversária discriminadora (7). Adicionalmente, a rede neuronal do tipo generativa adversária geradora (6) compreende um módulo de simulação de ruído de uma transmissão física (18) no sistema de codificação (1), que irá, no processo de treino das redes neuronais, adicionar ruído nas imagens de entrada, sendo o treino executado para gerar e identificar a informação escondida nestas condições.

EFEITOS VANTAJOSOS DA INVENÇÃO

A presente invenção apresenta elevada versatilidade no campo técnico dos documentos de segurança,

podendo ser aplicada para codificar e esconder informação numa imagem capturada, por exemplo o retrato de um indivíduo, em que a captura da imagem inclui uma fotografia, um fotograma, um vídeo, ou elementos de correlacionados, ocorrendo posteriormente a impressão em suporte físico ou eletrónico da imagem codificada, a fim de produzir um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21).

A presente invenção aplica-se à codificação de imagens em geral, sendo particularmente útil para esconder uma mensagem secreta em imagens faciais, também designadas por retratos, no contexto dos documentos de segurança com uma imagem codificada por esteganografia resistente à impressão (21) enquadrados como documentos de identificação civil e documentos de viagem com leitura ótica (ID-MRTD). Nesse sentido, tais documentos podem ser os cartões de identidade nacionais ou os passaportes, entre outros documentos. Contudo, a presente invenção não se limita a documentos físicos, mas também inclui documentos digitais ou qualquer modelo de identidade de pessoas que use a biometria facial como elemento de segurança para verificação de identidade (1-para-1) e a identificação para autenticação de pessoas (1-para-muitos).

Um aspeto importante da presente invenção refere-se ao objetivo global, que permanece em toda a fase de treino das redes neuronais, de manter a perceção da imagem codificada inalterada ou com uma alteração mínima, quase impercetível, em relação à imagem original de uma face, por exemplo. Efetivamente, as imagens codificadas de acordo com a presente invenção mantêm praticamente as mesmas propriedades de perceção do que as imagens originais, sendo o processo otimizado de forma que esta alteração seja a menor possível, a fim de não alterar a capacidade de reconhecimento

facial de sistemas de verificação (1-para-1) e identificação facial (1-para-muitos). Esta característica é uma das principais vantagens da presente invenção em relação ao conhecido no estado da técnica.

A presente invenção também apresenta diversas vantagens em relação ao método StegaStamp, o qual será referenciado em algumas ocasiões ao longo da descrição pormenorizada. Uma primeira vantagem da presente invenção está relacionada com a sua adequação para o processamento de imagens pequenas com as dimensões especificadas para as imagens de documentos ID-MRTD. Conforme será visto posteriormente, para que se tenha essa funcionalidade, é simulado o ruído referente ao redimensionamento das imagens, ao passo que o StegaStamp só funciona de forma aceitável para imagens de maiores dimensões, o que torna a sua aplicação em documentos ID-MRTD impossível. Uma segunda vantagem da presente invenção está relacionada com a sua adequação para imagens com sobreposição de camadas translúcidas com hologramas ou outros elementos. Nesse sentido, é simulado o ruído referente à captação da imagem de faces (ou outras), cuja personalização é feita por baixo de camadas de segurança que podem ser translúcidas, mas causam distorção na imagem captada. Por outro lado, o StegaStamp não tem esta simulação de ruído, pelo que não resiste a esta forte distorção da imagem a ser validada. Uma terceira vantagem da presente invenção está relacionada com a preservação das propriedades de perceção de uma face humana, quer por seres humanos, quer por sistemas automáticos de reconhecimento facial, provocando alterações mínimas à estrutura da face, ao passo que o StegaStamp altera consideravelmente as faces humanas, provocando distorções que tornam a sua aplicação a documentos ID-MTD impossível.

BREVE DESCRIÇÃO DAS FIGURAS

Com o propósito de promover um entendimento dos princípios de acordo com as modalidades da presente invenção, será efetuada referência às modalidades ilustradas nas figuras e à linguagem utilizada para descrevê-las. De qualquer modo, deve ser entendido que não há intenção de limitar o âmbito da presente invenção ao conteúdo das figuras. Quaisquer alterações ou modificações posteriores das características inventivas aqui ilustradas e quaisquer aplicações adicionais dos princípios e modalidades da invenção ilustrados, que ocorreriam normalmente para um perito na especialidade tendo a posse desta descrição, estão considerados no âmbito da invenção reivindicada.

Figura 1 – ilustra uma modalidade do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão;
Figura 2 – ilustra uma modalidade do subsistema de treino incluindo uma rede neuronal do tipo generativa adversária;
Figura 3 – ilustra uma modalidade do sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão.

DESCRIÇÃO DOS MODOS DE REALIZAÇÃO

A presente invenção diz respeito, num primeiro aspeto, a um sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) compreendendo:

um módulo codificador para correção de erros

binário e introdução de redundância (9), o qual é configurado para converter uma mensagem secreta (8) numa mensagem binária a codificar (10); e

um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12), o qual é configurado para processar e selecionar uma imagem parcial a codificar (13) a partir de uma imagem integral a codificar (11); e

um módulo codificador (14), o qual é configurado para codificar a mensagem binária a codificar (10) na imagem parcial a codificar (13), resultando numa imagem parcial codificada (15); e

um módulo integrador de imagens (16), o qual é configurado para integrar a imagem parcial codificada (15) na imagem integral a codificar (11), resultando numa imagem integral codificada (17);

em que o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) é treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3).

A presente invenção diz respeito, num segundo aspeto, a um sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) compreendendo:

um subsistema de descodificação (4), o qual inclui um segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27), um módulo descodificador (29) e um módulo de correção de erros binários com redundância (31); e

um subsistema de validação da integridade (5), o qual inclui um módulo de validação de uma mensagem descodificada (33) e um módulo de validação da integridade do documento de segurança (36);

em que o segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) é configurado para processar e selecionar uma imagem parcial a validar (28) a partir de uma imagem integral a validar (26);

em que o módulo descodificador (29) é configurado para descodificar a imagem parcial a validar (28), resultando numa mensagem binária a validar (30);

em que o módulo de correção de erros binários com redundância (31) é configurado para converter a mensagem binária a validar (30) numa mensagem descodificada a validar (32);

em que o módulo de validação de uma mensagem descodificada (33) é configurado para analisar se a mensagem descodificada a validar (32) é correta e conclusiva;

em que o módulo de validação da integridade do documento de segurança (36) é configurado para analisar se uma mensagem secreta correta (34) confere autenticidade ao documento de segurança a validar (22).

A presente invenção diz respeito, num terceiro aspeto, a um método implementado por computador para codificar um documento de segurança com uma imagem codificada por esteganografia resistente à impressão compreendendo as seguintes etapas:

a) Conversão de uma mensagem secreta (8) numa mensagem binária a codificar (10) através de um módulo codificador para correção de erros binário e introdução de redundância (9);

b) Processamento e seleção de uma imagem parcial a codificar (13) a partir de uma imagem integral a

codificar (11) através de um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12);

c) Codificação da mensagem binária a codificar (10) na imagem parcial a codificar (13) através de um módulo codificador (14), resultando numa imagem parcial codificada (15);

d) Integração da imagem parcial codificada (15) na imagem integral a codificar (11) através de um módulo integrador de imagens (16), resultando numa imagem integral codificada (17);

em que o módulo codificador para correção de erros binário e introdução de redundância (9), o primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12), o módulo codificador (14) e o módulo integrador de imagens (16) são incorporados num sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1);

em que o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) ser treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3).

A presente invenção diz respeito, num quarto aspeto, a um método implementado por computador para a descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão compreendendo as seguintes etapas:

a) Processamento e seleção de uma imagem integral a validar (26) através de um segundo módulo de deteção facial, alinhamento, recorte e

redimensionamento (27), resultando numa imagem parcial a validar (28);

b) Descodificação da imagem parcial a validar (28) através de um módulo descodificador (29), resultando numa mensagem binária a validar (30);

c) Conversão da mensagem binária a validar (30) numa mensagem descodificada a validar (32) através de um módulo de correção de erros binários com redundância (31);

d) Análise da mensagem descodificada a validar (32) por um módulo de validação de uma mensagem descodificada (33), resultando na decisão se a mensagem descodificada a validar (32) é uma mensagem secreta correta (34) ou uma mensagem secreta incorreta (35);

e) Análise da mensagem secreta correta (34) por um módulo de validação da integridade do documento de segurança (36), resultando na decisão se o documento de segurança a validar (22) é autêntico.

A presente invenção diz respeito, num quinto aspeto, a um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21), compreendendo pelo menos uma imagem integral codificada (17) transmitida fisicamente num suporte físico ou num suporte eletrónico e ser preparado por meio do método definido conforme o terceiro aspeto da invenção.

A presente invenção diz respeito, num sexto aspeto, a um aparelho computacional compreendendo meios adaptados para executar as etapas do método definido conforme o terceiro aspeto da invenção.

A presente invenção diz respeito, num sétimo aspeto, a um aparelho computacional compreendendo meios adaptados para executar as etapas do método definido conforme o quarto aspeto da invenção.

A presente invenção diz respeito, num oitavo aspeto, a um programa de computador compreendendo instruções para propiciar que o aparelho computacional, conforme definido no sexto aspeto da invenção, execute as etapas do método definido conforme o terceiro aspeto da invenção.

A presente invenção diz respeito, num nono aspeto, a um programa de computador compreendendo instruções para propiciar que o aparelho computacional, conforme definido no sétimo aspeto da invenção, execute as etapas do método definido conforme o quarto aspeto da invenção.

A presente invenção diz respeito, num décimo aspeto, a um meio de leitura por um aparelho computacional compreendendo a instalação do programa de computador, conforme definido no oitavo aspeto da invenção.

A presente invenção diz respeito, num décimo-primeiro aspeto, a um meio de leitura por um aparelho computacional compreendendo a instalação do programa de computador, conforme definido no nono aspeto da invenção.

Sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão

Do ponto de vista do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), é possível distinguir duas fases de aplicação da tecnologia da presente invenção: uma fase de treino e uma fase de teste. Como em qualquer sistema de aprendizagem máquina, em inglês *machine learning*, a fase de treino precede a fase de teste e destina-

se a afinar os parâmetros da rede de teste usando bancos de dados, em inglês *datasets*, relativamente grandes, em que a quantidade de dados de treino necessários depende do número de parâmetros e complexidade das redes a treinar.

Após a fase de treino, ficam operacionais na fase de teste o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) e o sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), o primeiro usado no momento da criação da imagem com uma mensagem secreta e também no controlo da qualidade de um documento de segurança na sua fase de produção e o segundo usado na aplicação de validação da integridade de um documento de segurança. Em termos de controlo da qualidade, cumpre destacar também o papel do sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), pois após a produção de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21) é efetuado o controlo da qualidade com base nos passos de descodificação, a fim de garantir que o documento de segurança produzido possa ser futuramente efetivamente autenticado conforme o método e o sistema de descodificação da presente invenção. Estes sistemas de codificação e descodificação usam ainda módulos de deteção facial, alinhamento e recorte da imagem e módulos de correção de erros binários com introdução de redundâncias.

Na figura 1 é ilustrada uma modalidade do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão. Em termos funcionais do sistema de codificação de um documento

de segurança com uma imagem codificada por esteganografia resistente à impressão (1), ocorre inicialmente a entrada de uma imagem integral a codificar (11), por exemplo uma imagem facial, e uma mensagem secreta (8). No modelo da presente invenção, a informação escondida será localizada numa parte específica da imagem integral a codificar (11). A parte pretendida da imagem integral a codificar (11) é, então, detetada e recortada, em inglês *crop*, por um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12), o qual utiliza por exemplo, os modelos BlazeFace, descrito em Valentin Bazarevsky, Yury Kartynnik, Andrey Vakunov, Karthik Raveendran, and Matthias Grundmann. Blazeface: Sub-millisecond neural face detection on mobile gpus. arXiv preprint arXiv:1907.05047, 2019) ou FaceNet, resultando numa imagem parcial a codificar (13). O primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12) tem a funcionalidade de padronizar as dimensões das imagens processadas para uma determinada dimensão compatível com as especificações exigidas pelo módulo subsequente. Nas modalidades preferidas da presente invenção, um primeiro módulo de pré-processamento de imagens (48) é incorporado antes do primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12) para compensar as imagens em termos de iluminação, por exemplo, ou em termos de cor, dimensão, estrutura, razão de aspeto ou algumas distorções (radial ou outra).

Paralelamente, conforme visualizado na figura 1, a mensagem secreta (8) é codificada por um módulo codificador para correção de erros binário e introdução de redundância (9), usando, por exemplo, os algoritmos de Reed-Solomon, descritos em Daniel Bleichenbacher, Aggelos Kiayias, and Moti Yung. Decoding of interleaved reed solomon codes over noisy data. pages 97–108, 2003; e em Stephen B Wicker and

Vijay K Bhargava. An introduction to reed-solomon codes. Reed-Solomon codes and their applications, pages 1-16, 1994; ou Bose-Chaudhuri-Hocquenghem (BCH), descrito em George Forney. On decoding bch codes. IEEE Transactions on information theory, 11(4):549-557, 1965. A introdução de redundância na informação é fundamental para melhorar a taxa de sucesso do método de descodificação, executado pelo sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2).

Em seguida, é efetuada a codificação da mensagem binária a codificar (10) na imagem parcial a codificar (13) através de um módulo codificador (14), resultando numa imagem parcial codificada (15) com uma mensagem secreta escondida. O passo subsequente envolve a integração da imagem parcial codificada (15) na imagem integral a codificar (11) através de um módulo integrador de imagens (16), resultando numa imagem integral codificada (17). A referida integração pode ser feita, por exemplo, através da substituição da porção original da imagem integral pela porção da imagem parcial codificada.

Conforme ilustrado na figura 1, nas modalidades preferidas da presente invenção, o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) compreende um módulo de transmissão física de uma imagem integral codificada num documento de segurança (20) para um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21).

O primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12) contribui para um processo robusto de verificação de imagens de documentos ocultando uma mensagem secreta na imagem facial ou retrato. Dessa

forma, é necessário um modelo de deteção facial que propicie inicialmente a identificação da porção da imagem onde está presente o rosto da pessoa, em seguida que possa selecionar a região de interesse do rosto, a fim de descobrir a parte do rosto onde a mensagem secreta está escondida. É importante notar que o modelo de deteção facial deve revelar exatamente a parte do rosto com as informações codificadas. Para tal, a presente invenção aplica um qualquer método do estado da técnica, nomeadamente classificadores em cascata como o HAAR ou LBP, BlazeFace, MobileNets V2 float32, MobileNets V2 int8, SSD int8 MTCNN, CASCATA LBP (opencv) ou PRnet.

O OpenCV é um exemplo de aplicação abertas ao uso público que possui um conjunto significativo de ferramentas para detetar faces e características especiais em imagens.

Testes exaustivos permitem concluir que os métodos de deteção de face baseados em aprendizagem profunda, em inglês *deep learning*, são mais adequados, sendo ainda de considerar que o foco de utilização destes sistemas é para dispositivos móveis pelo que convém ter em linha de conta a capacidade computacional exigida pelos métodos. Desta forma, os métodos BlazeFace e o Mobilenet V1/V2 são arquiteturas de aprendizagem profunda significativamente mais rápidas e suficientemente precisas para dispositivos móveis modernos.

Adicionalmente, o método PRnet fornece uma solução completa para a deteção facial e análise de pose facial que aumenta a precisão de deteção sob variação de pose e oclusão. Sem perda de generalidade, nas modalidades preferenciais de acordo com a presente invenção, o PRnet é o método com o melhor desempenho para o propósito da presente invenção. De qualquer modo, qualquer módulo de deteção facial, alinhamento e recorte, desenvolvido no futuro com métodos mais sofisticados, pode ser utilizado no âmbito da presente invenção.

O módulo codificador para correção de erros binário e introdução de redundância (9) executa a codificação de uma mensagem, usando por exemplo o método BCH ou o método Reed-Solomon, e possui a função adicional de melhorar a estabilidade e a exatidão na descodificação. Este módulo permitirá que o módulo descodificador (29) detete e corrija, quando possível, erros na mensagem a partir da redundância da informação introduzida no módulo codificador (14). Desta forma, aumenta-se drasticamente o rácio de imagens que se conseguem descodificar com sucesso. Embora a escolha exata dos métodos de correção de erros e redundância não seja objeto da presente invenção, consideramos que os métodos atualmente conhecidos no estado da técnica mais adequados são os códigos BCH e os códigos Reed-Solomon.

O módulo codificador (14) é o núcleo do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1). O objetivo central do módulo codificador (14) é a geração de uma imagem realista que otimize dois objetivos concorrentes: por um lado, a capacidade do módulo descodificador (29) extrair a mensagem secreta da imagem codificada e, por outro lado, a preservação da estrutura da imagem e das propriedades de perceção da imagem por seres humanos e por sistemas automáticos de reconhecimento de imagens.

A arquitetura do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) baseia-se na conhecida estrutura de rede em U, também designada por UNets, com remoção das camadas de agrupamento, em inglês designadas por *pooling layers*, para manter a informação das mensagens secretas, que podem ser perdidas durante a fase de treino. Assim, por exemplo, o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia

resistente à impressão (1) recebe como entrada a imagem de um rosto alinhado e uma mensagem secreta (8), gerando uma imagem codificada com substancialmente o mesmo tamanho. É de notar que para efeitos de treino da rede codificadora, a mensagem secreta de treino é aleatória. Preferivelmente, a mensagem binária secreta pode ser transformada para coincidir com o tamanho da entrada no codificador.

Uma vez que o módulo codificador (14) não tem camadas de agrupamento, isto é *pooling layers*, é necessário que o desenho da arquitetura seja cuidadoso, combinando manualmente os parâmetros das convoluções de forma a evitar erros de ligação de camadas.

Nas modalidades preferidas de acordo com a presente invenção, a imagem integral a codificar (11) é um retrato facial e o documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21) é selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

Preferivelmente, o método implementado por computador para codificar um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, conforme o terceiro aspeto da invenção é executado pelo sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), conforme definido no primeiro aspeto da invenção.

Subsistema de treino incluindo uma rede neuronal do tipo generativa adversária

Nas modalidades preferidas do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), o

sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) é treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3), o qual compreende uma rede neuronal do tipo generativa adversária geradora (6) e uma rede neuronal do tipo generativa adversária discriminadora (7). A rede neuronal do tipo generativa adversária geradora (6) e a rede neuronal do tipo generativa adversária discriminadora (7) são executadas ao longo da fase de treino para ocultar e ler mensagens secretas em imagens, por exemplo imagens faciais e, por outro lado, são adicionadas várias camadas de simulação de ruído entre os módulos codificador e descodificador, de forma a criar imagens realistas durante o treino, com particular ênfase nos ruídos que simulam a transmissão por meio físico, como a impressão ou a transmissão para um ecrã.

Um módulo de simulação de ruído de uma transmissão física (18) é incorporado na rede neuronal do tipo generativa adversária geradora (6) com a função de simular o ruído que ocorre numa transmissão física de uma imagem codificada, nomeadamente a impressão ou visualização num ecrã e subsequente captura de imagem por câmaras digitais, ou câmaras analógicas com subsequente digitalização. Assim, são aplicados às imagens de treino (40) vários tipos de ruído, resultando em imagens de treino com ruído simulado (19), em que o treino iterativo contribui para a robustez do método de verificação da autenticidade de um documento de segurança a validar na fase de teste do processamento pelo sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2).

Os métodos do estado da técnica no campo da simulação de ruído são os métodos HiDDeN, descrito em: ZHU,

Jiren, et al. Hidden: Hiding data with deep networks. In: Proceedings of the European conference on computer vision (ECCV). 2018. p. 657-672; e StegaStamp, descrito em: Matthew Tancik, Ben Mildenhall, Ren Ng Invisible Hyperlinks in Physical Photographs; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2117-2126. O sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) da presente invenção engloba o conjunto dos mesmos ruídos aplicados por estes métodos, tais como a distorção geométrica, incluindo a distorção de perspetiva; o desfoque ótico e de movimento; o ruído da câmara; a manipulação de cores e a compressão de imagem, por exemplo pelo formato JPEG. A distorção de perspetiva pode ser simulada por uma homografia aleatória que simula o efeito da câmara que não está alinhada com precisão com o plano da imagem codificada. O desfoque ótico, em inglês *defocus*, e o desfoque por movimento, em inglês *motion blur*, podem ser resultantes tanto do movimento da câmara como do foco automático impreciso, que é muito comum em dispositivos móveis. Para simular o desfoque de movimento, é aplicado um núcleo linear de desfoque com ângulo aleatório e largura entre 3 e 7 pixéis. O ruído da câmara, que inclui ruído fotónico, ruído escuro e ruído de disparo do sistema de câmaras estão bem descritos e documentados pelos trabalhos citados do estado da arte. A manipulação de cores, que é um ruído que resulta, quer de impressoras, quer de ecrãs, numa gama limitada em comparação com o espaço de cores RGB completo e inclui mudança de tonalidade, saturação, brilho e contraste. A presente invenção também engloba um ruído referente a uma compressão de imagem, por exemplo no formato JPEG, que afeta a imagem quando é armazenada num formato com perda, como é o caso do JPEG.

Com relação aos ruídos de transmissão física no contexto dos documentos ID-MRTD, também são englobados na presente invenção os ruídos de sobreposição por material sobreposto translúcido, os quais são pertinentes quando sobre as imagens dos documentos de segurança é colocada um holograma transparente ou translúcido ou outro material em camadas e que provoca uma degradação na perceção de uma imagem, por exemplo uma imagem da face. As imagens das faces são personalizadas nos documentos de segurança em camadas abaixo de camadas, denominadas por *layers* na língua inglesa, que lhe estão acima, visto que a construção dos próprios documentos se faz em camadas, algumas delas com elementos de segurança, como é o caso do holograma transparente. Assim, a imagem captada de uma face num documento destes tem sempre parte da imagem obstruída por elementos semitransparentes ou translúcidos.

Todos estes ruídos são apropriados para a integração entre a rede neuronal do tipo generativa adversária geradora (6) e a rede neuronal do tipo generativa adversária discriminadora (7) na fase de treino, nomeadamente com a introdução do ruído entre o módulo codificador de treino (41) e o módulo descodificador de treino (42) na rede neuronal do tipo generativa adversária geradora (6). Todos as componentes de ruído adicionadas têm um parâmetro escalar que governa a intensidade da distorção, para efeitos de treino da rede.

Contudo, estes ruídos apresentados, existentes nos modelos do estado da técnica citados, não são por si suficientes para cumprir um dos objetivos principais do problema a resolver pela presente invenção, nomeadamente a necessidade de manter inalterada, ou com uma alteração mínima, praticamente impercetível, a estrutura facial da imagem codificada, isto é, manter quase inalteradas as

capacidades de perceção facial da imagem por um ser humano ou por um sistema de reconhecimento facial, quer por verificação (1-para-1) quer por identificação (1-para-muitos). Para tal, um módulo de redimensionamento de uma imagem capturada (25) é incorporado na rede neuronal do tipo generativa adversária geradora (6), sendo que o módulo de redimensionamento de uma imagem capturada (25) introduz ruído ao redimensionar as imagens das faces para imagens mais pequenas, que consequentemente reduzem drasticamente a resolução da imagem facial e, por isso, permitem treinar a rede para que a descodificação seja possível em imagens de faces menores.

O módulo de redimensionamento de uma imagem capturada (25) contribui para que o sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) seja capaz de ler uma mensagem a partir de uma pequena imagem facial impressa em documentos. A dimensão da imagem é um ponto fundamental da presente invenção, visto que conforme se diminui o tamanho de uma imagem facial, por exemplo, maior é a dificuldade em codificar e descodificar com sucesso uma mensagem secreta. Existem documentos ID-MRTD com inúmeras dimensões de imagem facial, contudo, o estado da técnica é determinado a partir de algumas instituições internacionais mais influentes e regem-se, geralmente, por padrões internacionais efetivos ou de facto. Alguns dos documentos mais importantes para esta definição da dimensão da imagem são o documento ICAO 9303 e a norma internacional ISO IEC 19794:5.

Apesar de não existir uma dimensão específica para a qual é pertinente resolver o problema que a presente invenção resolve, a prática determina que as imagens dos documentos de segurança a validar (22) têm usualmente uma

largura no intervalo de cerca de 2 cm a cerca de 4 cm.

Consequentemente, um contributo técnico adicional apresentado pela presente invenção refere-se à incorporação do módulo de redimensionamento de uma imagem capturada (25), cujo objetivo é redimensionar as imagens codificadas de entrada para dimensões próprias dos documentos ID-MRTD. Esta transformação da imagem representa um nível de ruído adicional, cujo treino permite melhorar substancialmente a capacidade de o módulo descodificador (29) recuperar a mensagem secreta com sucesso em retratos de menores dimensões.

Ainda no que toca à rede neuronal do tipo generativa adversária discriminadora (7) que engloba o módulo de cálculo de uma função de perda (45), esta consiste num conjunto de componentes que permitem durante o treino otimizar os parâmetros das várias redes, sendo utilizadas componentes de perda padrão para este tipo de redes e, adicionalmente, uma função de perda cujo efeito é preservar a aparência do imagem codificada durante o treino.

A rede neuronal do tipo generativa adversária discriminadora (7) inclui um módulo de deteção facial, alinhamento e recorte de treino (44), uma CNN e uma camada simples densa, em inglês *Simple Dense layer*, que é uma rede de avanço rápido, em inglês *fast forward network*. O módulo de deteção facial, alinhamento e recorte de treino (44) recebe a imagem codificada e recorta a parte apropriada para a entrada no módulo de cálculo de uma função de perda (45). Nas modalidades preferidas, a rede neuronal do tipo generativa adversária discriminadora (7) compreende uma função de perda percetual LPIPLS, sigla da língua inglesa *Learning Perceptual Image Patch Similarity*, e uma função de perda crítica para imagens codificadas, uma camada densa e uma função de perda binária de entropia cruzada, em inglês

cross entropy, para mensagens descodificadas e uma componente de descrição facial, em inglês *face embedding*. A função de perda percetual LPIPLS é descrita em: Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 586-595).

Nas modalidades preferidas, o módulo de cálculo de uma função de perda (45) inclui a função de perda percetual LPIPLS e o componente de descrição facial, em inglês *face embedding*.

No que se refere à função de perda percetual LPIPLS, esta demonstrou a sua eficácia em modelos do estado da técnica, tais como no modelo StegaStamp. Portanto, no modelo da presente invenção é usada como função de perda percetual.

No que se refere ao componente de descrição facial, em inglês *face embedding*, considerando um dos objetivos principais do método da presente invenção a preservação da estrutura facial e da sua representação de alto nível, a função de perda foi modificada para incluir a função de semelhança que é estimada pela saída de um modelo FaceNet. O modelo FaceNet utiliza a arquitetura Inception Resnet V1 que foi treinada com o conjunto de dados, em inglês *dataset*, VGG2 e é descrito em: Florian Schroff, Dmitry Kalenichenko, James Philbin; FaceNet: A Unified Embedding for Face Recognition and Clustering; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815-823. Este modelo recebe uma imagem facial RGB de 160x160 pixéis e disponibiliza à saída um vetor de características faciais de 128 dimensões. A distância Euclidiana entre os dois conjuntos de características expressa a dissemelhança entre as duas imagens de origem. O

modelo da presente invenção calcula assim a distância Euclidiana para minimizar as características faciais entre as imagens original e codificada, durante o processo de treino.

Aos dois termos particularmente preferidos da função de perda, nomeadamente a função de perda percetual LPIPLS e o componente de descrição facial, em inglês *face embedding*, são adicionados outros termos de perda que são usados nos modelos do estado da técnica, acima mencionados, concretamente os modelos Stegastamp e Hidden.

Dessa forma, o módulo de cálculo de uma função de perda (45) inclui pelo menos um dos seguintes componentes: função de perda percetual LPIPS - L_P; função de perda FaceNet - L_F, que se refere a um sistema de reconhecimento biométrico arbitrário que mede a proximidade de dois modelos faciais pela distância Euclidiana; função de perda de Wasserstein - L_W, que é utilizada como perda percetual para o pipeline codificador/descodificador; função de perda de regularização residual - L_R; e função de perda de entropia cruzada - L_B, que treina a rede de descodificador para recuperar a mensagem.

Dessa forma, a fase de treino engloba os principais módulos do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) e do sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) e vai permitir a afinação dos parâmetros dos módulos de codificação e descodificação que são usados na fase de teste.

Nas modalidades particularmente preferidas da invenção, durante a execução do subsistema de treino incluindo uma rede neuronal do tipo generativa adversária

(3), a função de perda global é, então, a soma ponderada de cinco termos de perda, ou seja, dado pela equação (Loss = F*L_F + P*L_P + W*L_W + R*L_R + B*L_B), onde F, P, W, R e B são os pesos escalares das componentes da função de perda. Na fase inicial do treino, F, P, W e R estão inicialmente definidos a zero e B é definido para o valor 0,01, até que o descodificador atinja alta precisão com estes pesos. Nas experiências efetuadas com a presente invenção, geralmente são necessários cerca de 500 a cerca de 700 passos para atingir a alta precisão. Depois de atingida alta precisão, estes pesos podem ser aumentados linearmente a cada passo. Com isso, é possível melhorar lentamente a eficácia da função de perda global, aumentando estes coeficientes escalares.

Conforme ilustrado na figura 2, a fim de representar as modalidades preferidas da invenção durante a fase de treino, a rede neuronal do tipo generativa adversária geradora (6) é configurada para processar uma mensagem secreta de treino (39) e uma imagem de treino (40) por intermédio de um módulo codificador de treino (41), um módulo de simulação de ruído de uma transmissão física (18), um módulo de redimensionamento de uma imagem capturada (25) e um módulo descodificador de treino (42), resultando numa imagem artificial com uma mensagem codificada de treino (43). Em seguida e em paralelo, a rede neuronal do tipo generativa adversária discriminadora (7) é configurada para processar a imagem artificial com uma mensagem codificada de treino (43) por intermédio de um módulo de deteção facial, alinhamento e recorte de treino (44) e um módulo de cálculo de uma função de perda (45), resultando numa mensagem secreta recuperada (46); em que o treino é executado através de uma pluralidade de passos de processamento de mensagens secretas de treino (39) e imagens de treino (40) e geração de mensagens secretas recuperadas (46), em que o processo de

treino da rede neuronal do tipo generativa adversária geradora (6) ocorre por intermédio de um processo de propagação do erro, calculado pelo módulo de cálculo de uma função de perda (45), e de um processo de ajuste dos parâmetros da rede geradora, executado por um módulo de ajuste dos parâmetros da rede geradora (47), nomeadamente os parâmetros de funcionamento do módulo codificador de treino (41) e do módulo descodificador de treino (42).

Na transição da fase de treino pela rede neuronal do tipo generativa adversária para a fase de teste do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) e do sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), o módulo codificador (14) é baseado no módulo codificador de treino (41); o módulo descodificador (29) é baseado no módulo descodificador de treino (42); e o primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12) e o segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) são baseados no módulo de deteção facial, alinhamento e recorte de treino (44). Nos modos de concretização particularmente preferidos, o primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12) e o segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) podem ser essencialmente o mesmo módulo de execução.

Sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão

A fim de verificar se um documento de segurança a

validar (22) é autêntico e conforme ilustrado na figura 3, um sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) opera por intermédio de uma etapa inicial de processamento e seleção de uma imagem integral a validar (26) através de um segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27), resultando numa imagem parcial a validar (28). O segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) é responsável, por exemplo, pela deteção da face principal numa imagem integral a validar (26) de entrada e, posteriormente, deteta a parte codificada da imagem facial. Adicionalmente, o segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) tem a funcionalidade de padronizar as dimensões das imagens processadas para uma determinada dimensão compatível com as especificações exigidas pelo módulo subsequente.

Conforme ilustrado na figura 3, nas modalidades preferidas da presente invenção, um segundo módulo de pré-processamento de imagens (49) é incorporado antes do segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) para compensar as imagens em termos de iluminação, por exemplo, ou em termos de cor, dimensão, estrutura, razão de aspeto ou algumas distorções (radial ou outra).

Noutras modalidades preferidas da invenção, a imagem integral capturada a validar (24) é obtida a partir da captura de uma imagem de segurança incluída no documento de segurança a validar (22) através de um módulo de captura digital de uma imagem integral incluída em pelo menos um documento de segurança a validar (23), o qual é incorporado numa câmara digital ou faz parte da integração entre uma

câmara analógica cuja imagem é depois digitalizada por um dispositivo de digitalização configurado para o efeito. Ainda mais preferivelmente, a imagem integral capturada a validar (24) ser um retrato facial e o documento de segurança a validar (22) ser selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

Dessa forma, o segundo módulo de pré-processamento de imagens (49) pode receber para processamento tanto uma imagem integral a validar (26), quanto uma imagem integral capturada a validar (24), consideradas equivalentes no contexto da presente invenção.

Retomando a ilustração da figura 3, a etapa seguinte inclui a descodificação da imagem parcial a validar (28) através de um módulo descodificador (29), resultando numa mensagem binária a validar (30). O próximo passo inclui a conversão da mensagem binária a validar (30) numa mensagem descodificada a validar (32), composta num número ou cadeia de caracteres, em inglês designada por *string*, através de um módulo de correção de erros binários com redundância (31).

Para o módulo descodificador (29), são aplicadas uma rede neuronal convolucional (CNN) juntamente com uma Rede Especial Transformadora ou STN, em inglês *Special Transformer Network*, preferencialmente baseada nos modelos StegaStamp e HiDDeN. A rede STN ajuda a isolar a região apropriada e a normalizar a sua escala, o que pode simplificar a tarefa de descodificação da informação escondida por esteganografia e levar a um melhor desempenho. A rede remove a parte espacialmente invariante da imagem codificada, aplicando uma transformação afim aprendida, seguida por uma operação de interpolação. Nas modalidades preferenciais da invenção, a rede STN é colocada antes da rede CNN.

Ainda no contexto da figura 3, após a execução pelo subsistema de descodificação (4), parte-se para o processamento pelo subsistema de validação da integridade (5) do documento a validar, em que é efetuado um primeiro passo de análise da mensagem descodificada a validar (32) por um módulo de validação de uma mensagem descodificada (33), resultando na decisão se a mensagem descodificada a validar (32) é uma mensagem secreta correta (34) ou uma mensagem secreta incorreta (35). Este passo é executado por intermédio de algoritmos de validação, usualmente através de uma função de hash ou algoritmo de verificação através de um *checksum*. No final deste módulo o subsistema de validação da integridade (5) tem à saída a recuperação de uma mensagem secreta correta (34) ou a informação de que a descodificação foi inconclusiva, tendo em vista a recuperação de uma mensagem secreta incorreta (35). Em seguida, é efetuada uma etapa de análise da mensagem secreta correta (34) por um módulo de validação da integridade do documento de segurança (36), resultando na decisão se o documento de segurança a validar (22) é autêntico.

Nas modalidades particularmente preferidas da presente invenção, o sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) apresenta um módulo descodificador (29), que inclui pelo menos uma rede especial transformadora e pelo menos uma rede neuronal convolucional.

Em certos modos de realização da presente invenção, o sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) é integrado com o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia

resistente à impressão (1).

Preferencialmente, o método implementado por computador para a descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo com o quarto aspeto da invenção, é executado pelo sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), conforme o segundo aspeto da invenção.

O módulo de validação de uma mensagem descodificada (33) tem como objetivo analisar a informação extraída ou recuperada pelo módulo descodificador (29) de uma imagem parcial a validar (28) e verificar o cumprimento de um conjunto de regras de validação necessárias. A mensagem descodificada a validar (32) entra no módulo de validação de uma mensagem descodificada (33) depois de corrigidos os eventuais erros na descodificação, através dos métodos de correção de erros, como por exemplo os métodos BCH ou Reed-Solomon.

O módulo de validação de uma mensagem descodificada (33) aplica então um sistema de verificação à mensagem descodificada a validar (32) produzindo uma resposta binária do cumprimento dessa verificação: mensagem secreta correta (34) ou mensagem secreta incorreta (35). Embora o método concreto de validação da mensagem não esteja no âmbito da presente invenção, alguns dos principais métodos adequados a este módulo incluem o uso de uma função de hash criptográfica, que quando aplicada a uma parte da mensagem gera uma informação que deverá coincidir com a restante parte da mensagem, transmitidas em simultâneo, ou também o uso de dígitos verificadores, designados em inglês como *check digits* ou *checksums*, os quais aplicam um conjunto de cálculos

a uma parte da mensagem e cujo resultado deve coincidir com a restante parte da mensagem, transmitidas em simultâneo.

No caso do módulo de validação de uma mensagem descodificada (33) produzir uma mensagem secreta incorreta (35), a imagem integral capturada a validar (24) é considerada inconclusiva, uma vez que a incorreção da mensagem se deve, neste caso, a um erro de transmissão, isto é, a transmissão física provocou uma degradação elevada na imagem que tornou impossível recuperar a mensagem completamente.

Por outro lado, no caso do módulo de validação de uma mensagem descodificada (33) produzir uma mensagem secreta correta (34), a mensagem recuperada é disponibilizada à saída deste módulo de forma a que seja validada pelo módulo de validação da integridade do documento de segurança (36).

O módulo de validação da integridade do documento de segurança (36) tem como objetivo verificar se a mensagem secreta correta (34) é a mensagem secreta esperada e que garante a integridade do documento de segurança, por exemplo um ID-MRTD. É usual, embora não seja obrigatório, que a mensagem seja, ou pelo menos contenha, o número do documento ID-MRTD e possivelmente mais alguns dados pessoais do portador do documento. Nestes casos, o processo de validação de identidade compara a mensagem recuperada e validada com o número do documento e as demais informações contidas na mensagem secreta.

O resultado módulo de validação da integridade do documento de segurança (36) é uma resposta binária relativa à integridade do documento: uma decisão positiva sobre a autenticidade do documento de segurança (37) ou uma decisão negativa sobre a autenticidade do documento de segurança (38). O documento de segurança é considerado autêntico

(válido) se a mensagem secreta correta (34) é validada e igual à mensagem esperada, construída pelo sistema de validação. Nos restantes casos, o documento é considerado não autêntico (inválido).

Numa possibilidade alternativa da presente invenção, parte da informação necessária à tomada de decisão de integridade está armazenada numa base de dados segura remota. Nestes casos, a informação recuperada e validada da mensagem secreta correta (34), eventualmente complementada com informação adicional constante na imagem do documento ou introduzida diretamente no equipamento de validação, é usada para aceder a uma base de dados segura remota onde a tomada de decisão será feita. A informação recebida pelo sistema remoto e a informação armazenada nesse sistema e desbloqueada serão então submetidas ao módulo de validação da integridade do documento de segurança (36), de forma a tomar uma decisão relativa à integridade do documento. Esta decisão binária é devolvida ao equipamento de validação.

No caso desta alternativa remota, o módulo de validação da integridade do documento de segurança (36) pode produzir uma decisão inconclusiva, nas situações em que a ligação ao sistema remoto seja impossível ou possa estar comprometida em termos de conectividade ou segurança.

Como utilizados nesta descrição, as expressões “cerca de” e “aproximadamente” referem-se a um intervalo de valores de mais ou menos 10% o número especificado.

Como utilizado ao longo deste pedido de patente, o termo “ou” é utilizado no sentido inclusivo ao invés do sentido exclusivo, a menos que o sentido exclusivo seja claramente definido numa situação específica. Neste contexto, uma frase do tipo “X utiliza A ou B” deve ser interpretada como incluindo todas as combinações inclusivas

pertinentes, por exemplo “X utiliza A”, “X utiliza B” e “X utiliza A e B”.

Como utilizado ao longo deste pedido de patente, os artigos indefinidos “um” ou uma” devem ser interpretados geralmente como “um ou mais” e “uma ou mais”, a menos que o sentido de uma modalidade singular seja claramente definido numa situação específica.

Como apresentados nesta descrição, os termos relacionados com exemplos devem ser interpretados com o propósito de ilustrar um exemplo de algo e sem o propósito de indicar uma preferência.

Como utilizado nesta descrição, a expressão “substancialmente” significa que o valor real está dentro do intervalo de cerca de 10% do valor desejado, variável ou limite relacionado, particularmente dentro de cerca de 5% do valor desejado, variável ou limite relacionado ou especialmente dentro de cerca de 1% do valor desejado, variável ou limite relacionado.

A matéria-objeto descrita acima é fornecida como uma ilustração da presente invenção e não deve ser interpretada de modo a limitá-la. A terminologia utilizada com o propósito de descrever modalidades específicas, de acordo com a presente invenção, não deve ser interpretada para limitar a invenção. Como usados na descrição, os artigos definidos e indefinidos, na sua forma singular, visam a interpretação de incluírem também as formas plurais, a não ser que o contexto da descrição indique, explicitamente, o contrário. Será entendido que os termos “compreender” e “incluir”, quando usados nesta descrição, especificam a presença das características, dos elementos, dos componentes, das etapas e das operações relacionadas, mas não excluem a possibilidade de outras características, elementos, componentes, etapas e operações também estarem

contempladas.

Todas as alterações, desde que não modifiquem as características essenciais das reivindicações que se seguem, devem ser consideradas dentro do âmbito da proteção da presente invenção.

LISTA DE INDICAÇÕES DE REFERÊNCIA

1. Um sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão
2. Um sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão
3. Um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária
4. Um subsistema de descodificação
5. Um subsistema de validação da integridade
6. Uma rede neuronal do tipo generativa adversária geradora
7. Uma rede neuronal do tipo generativa adversária discriminadora
8. Uma mensagem secreta
9. Um módulo codificador para correção de erros binário e introdução de redundância
10. Uma mensagem binária a codificar
11. Uma imagem integral a codificar
12. Um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento
13. Uma imagem parcial a codificar
14. Um módulo codificador

15. Uma imagem parcial codificada

16. Um módulo integrador de imagens

17. Uma imagem integral codificada

18. Um módulo de simulação de ruído de uma transmissão física

19. Uma imagem de treino com ruído simulado

20. Um módulo de transmissão física de uma imagem integral codificada num documento de segurança

21. Um documento de segurança com uma imagem codificada por esteganografia resistente à impressão

22. Um documento de segurança a validar

23. Um módulo de captura digital de uma imagem integral incluída num documento de segurança a validar

24. Uma imagem integral capturada a validar

25. Um módulo de redimensionamento de uma imagem capturada

26. Uma imagem integral a validar

27. Um segundo módulo de deteção facial, alinhamento, recorte e redimensionamento

28. Uma imagem parcial a validar

29. Um módulo descodificador

30. Uma mensagem binária a validar

31. Um módulo de correção de erros binários com redundância

32. Uma mensagem descodificada a validar

33. Um módulo de validação de uma mensagem descodificada

34. Uma mensagem secreta correta

35. Uma mensagem secreta incorreta

36. Um módulo de validação da integridade do documento de segurança

37. Uma decisão positiva sobre a autenticidade do documento de segurança

38. Uma decisão negativa sobre a autenticidade do documento de segurança

39. Uma mensagem secreta de treino

40. Uma imagem de treino

41. Um módulo codificador de treino

42. Um módulo descodificador de treino

43. Uma imagem artificial com uma mensagem codificada de treino

44. Um módulo de deteção facial, alinhamento e recorte de treino

45. Um módulo de cálculo de uma função de perda

46. Uma mensagem secreta recuperada

47. Um módulo de ajuste de parâmetros da rede geradora

48. Um primeiro módulo de pré-processamento de imagens

49. Um segundo módulo de pré-processamento de imagens

LISTA DE CITAÇÕES

Valentin Bazarevsky, Yury Kartynnik, Andrey Vakunov, Karthik Raveendran, and Matthias Grundmann. Blazeface: Sub-millisecond neural face detection on mobile gpus. arXiv preprint arXiv:1907.05047, 2019;

Daniel Bleichenbacher, Aggelos Kiayias, and Moti Yung. ecoding of interleaved reed solomon codes over noisy data. pages 97–108, 2003;

Stephen B Wicker and Vijay K Bhargava. An

introduction to reed-solomon codes. Reed-Solomon codes and their applications, pages 1–16, 1994;

George Forney. On decoding bch codes. IEEE Transactions on information theory, 11(4):549–557, 1965;

ZHU, Jiren, et al. Hidden: Hiding data with deep networks. In: Proceedings of the European conference on computer vision (ECCV). 2018. p. 657-672;

Matthew Tancik, Ben Mildenhall, Ren Ng Invisible Hyperlinks in Physical Photographs; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2117-2126;

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 586-595);

Florian Schroff, Dmitry Kalenichenko, James Philbin; FaceNet: A Unified Embedding for Face Recognition and Clustering; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815-823.

REIVINDICAÇÕES

1. Um sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) compreendendo:

um módulo codificador para correção de erros binário e introdução de redundância (9), o qual é configurado para converter uma mensagem secreta (8) numa mensagem binária a codificar (10); e

um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12), o qual é configurado para processar e selecionar uma imagem parcial a codificar (13) a partir de uma imagem integral a codificar (11); e

um módulo codificador (14), o qual é configurado para codificar a mensagem binária a codificar (10) na imagem parcial a codificar (13), resultando numa imagem parcial codificada (15); e

um módulo integrador de imagens (16), o qual é configurado para integrar a imagem parcial codificada (15) na imagem integral a codificar (11), resultando numa imagem integral codificada (17);

em que o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) é treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3), **caracterizado por** o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) ser configurado para ser treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3), o qual compreende uma rede neuronal do tipo generativa adversária

geradora (6) e uma rede neuronal do tipo generativa adversária discriminadora (7); e em que
a rede neuronal do tipo generativa adversária geradora (6) é configurada para processar uma mensagem secreta de treino (39) e uma imagem de treino (40) por intermédio de um módulo codificador de treino (41), um módulo de simulação de ruído de uma transmissão física (18), um módulo de redimensionamento de uma imagem capturada (25) e um módulo descodificador de treino (42), resultando numa imagem artificial com uma mensagem codificada de treino (43); e por a rede neuronal do tipo generativa adversária discriminadora (7) ser configurada para processar a imagem artificial com uma mensagem codificada de treino (43) por intermédio de um módulo de deteção facial, alinhamento e recorte de treino (44) e um módulo de cálculo de uma função de perda (45), resultando numa mensagem secreta recuperada (46); em que o treino é executado através de uma pluralidade de passos de processamento de mensagens secretas de treino (39) e imagens de treino (40) e geração de mensagens secretas recuperadas (46) até que o módulo de cálculo de uma função de perda (45) calcule uma eficiência satisfatória na recuperação da mensagem secreta.

2. O sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), de acordo com a reivindicação anterior, **caracterizado por** compreender um módulo (20) de transmissão física da imagem integral codificada (17) para um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21).

3. O sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia

resistente à impressão (1), de acordo com qualquer uma das reivindicações anteriores, **caracterizado por** a imagem integral a codificar (11) ser um retrato facial e o documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21) ser selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

4. Um sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2) **caracterizado por** ser integrado com o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), conforme definido em qualquer uma das reivindicações anteriores, e **por** compreender:

um subsistema de descodificação (4), o qual inclui um segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27), um módulo descodificador (29) e um módulo de correção de erros binários com redundância (31); e

um subsistema de validação da integridade (5), o qual inclui um módulo de validação de uma mensagem descodificada (33) e um módulo de validação da integridade do documento de segurança (36);

em que o segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27) é configurado para processar e selecionar uma imagem parcial a validar (28) a partir de uma imagem integral a validar (26);

em que o módulo descodificador (29) é configurado para descodificar a imagem parcial a validar (28), resultando numa mensagem binária a validar (30);

em que o módulo de correção de erros binários com redundância

(31) é configurado para converter a mensagem binária a validar (30) numa mensagem descodificada a validar (32);
em que o módulo de validação de uma mensagem descodificada (33) é configurado para analisar se a mensagem descodificada a validar (32) é correta e conclusiva;
em que o módulo de validação da integridade do documento de segurança (36) é configurado para analisar se uma mensagem secreta correta (34) confere autenticidade ao documento de segurança a validar (22); e em que
o módulo de validação da integridade do documento de segurança (36) é configurado para produzir uma resposta binária relativa à integridade do documento: uma decisão positiva sobre a autenticidade do documento de segurança (37) ou uma decisão negativa sobre a autenticidade do documento de segurança (38); e em que o documento de segurança é considerado autêntico se a mensagem secreta correta (34) é validada e igual à mensagem secreta esperada.

5. O sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), de acordo com a reivindicação anterior, **caracterizado por** o módulo descodificador (29) incluir pelo menos uma rede especial transformadora e pelo menos uma rede neuronal convolucional.

6. O sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), de acordo com qualquer uma das reivindicações 4 e 5, **caracterizado por** incluir um módulo de captura digital de uma imagem integral incluída num documento de segurança a validar (23), o qual é configurado para capturar uma imagem

integral capturada a validar (24) de uma imagem de segurança incluída no documento de segurança a validar (22).

7. O sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), de acordo com qualquer uma das reivindicações 4 a 6, **caracterizado por** a imagem integral capturada a validar (24) ser um retrato facial e o documento de segurança a validar (22) ser selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

8. Um método implementado por computador para codificar um documento de segurança com uma imagem codificada por esteganografia resistente à impressão **caracterizado por** ser executado pelo sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1), conforme definido em qualquer uma das reivindicações 1 a 3, e **por** compreender as seguintes etapas:

a) Conversão de uma mensagem secreta (8) numa mensagem binária a codificar (10) através de um módulo codificador para correção de erros binário e introdução de redundância (9);

b) Processamento e seleção de uma imagem parcial a codificar (13) a partir de uma imagem integral a codificar (11) através de um primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12);

c) Codificação da mensagem binária a codificar (10) na imagem parcial a codificar (13) através de um módulo

codificador (14), resultando numa imagem parcial codificada (15);

d) Integração da imagem parcial codificada (15) na imagem integral a codificar (11) através de um módulo integrador de imagens (16), resultando numa imagem integral codificada (17);

em que o módulo codificador para correção de erros binários e introdução de redundância (9), o primeiro módulo de deteção facial, alinhamento, recorte e redimensionamento (12), o módulo codificador (14) e o módulo integrador de imagens (16) são incorporados num sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1);

em que o sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) é treinado por um subsistema de treino incluindo uma rede neuronal do tipo generativa adversária (3), o qual compreende uma rede neuronal do tipo generativa adversária geradora (6) e uma rede neuronal do tipo generativa adversária discriminadora (7); e em que

o treino do sistema de codificação de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (1) é efetuado através do processamento pela rede neuronal do tipo generativa adversária geradora (6) de uma mensagem secreta de treino (39) e uma imagem de treino (40) por intermédio de um módulo codificador de treino (41), um módulo de simulação de ruído de uma transmissão física (18), um módulo de redimensionamento de uma imagem capturada (25) e um módulo descodificador de treino (42), resultando numa imagem artificial com uma mensagem codificada de treino (43); e por a rede neuronal do tipo generativa adversária discriminadora (7) processar a imagem artificial com uma mensagem codificada

de treino (43) por intermédio de um módulo de deteção facial, alinhamento e recorte de treino (44) e um módulo de cálculo de uma função de perda (45), resultando numa mensagem secreta recuperada (46); em que o treino é executado através de uma pluralidade de passos de processamento de mensagens secretas de treino (39) e imagens de treino (40) e geração de mensagens secretas recuperadas (46) até que o módulo de cálculo de uma função de perda (45) calcule uma eficiência satisfatória na recuperação da mensagem secreta.

9. O método implementado por computador para codificar um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo com a reivindicação 8, **caracterizado por** incluir uma etapa adicional posterior de transmissão física da imagem integral codificada (17) para um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21), sendo esta etapa executada por um módulo de transmissão física de uma imagem integral codificada (20).

10. O método implementado por computador para codificar um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo com qualquer uma das reivindicações 8 e 9, **caracterizado por** a imagem integral a codificar (11) ser um retrato facial e o documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21) ser selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

11. Um método implementado por computador para a descodificação e validação da integridade de um documento de

segurança com uma imagem codificada por esteganografia resistente à impressão **caracterizado por** ser executado pelo sistema de descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (2), conforme definido em qualquer uma das reivindicações 4 a 7 e **por** compreender as seguintes etapas:

a) Processamento e seleção de uma imagem integral a validar (26) através de um segundo módulo de deteção facial, alinhamento, recorte e redimensionamento (27), resultando numa imagem parcial a validar (28);

b) Descodificação da imagem parcial a validar (28) através de um módulo descodificador (29), resultando numa mensagem binária a validar (30);

c) Conversão da mensagem binária a validar (30) numa mensagem descodificada a validar (32) através de um módulo de correção de erros binários com redundância (31);

d) Análise da mensagem descodificada a validar (32) por um módulo de validação de uma mensagem descodificada (33), resultando na decisão se a mensagem descodificada a validar (32) é uma mensagem secreta correta (34) ou uma mensagem secreta incorreta (35);

e) Análise da mensagem secreta correta (34) por um módulo de validação da integridade do documento de segurança (36), resultando na decisão se o documento de segurança a validar (22) é autêntico; e em que

o módulo de validação da integridade do documento de segurança (36) produz uma resposta binária relativa à integridade do documento: uma decisão positiva sobre a

autenticidade do documento de segurança (37) ou uma decisão negativa sobre a autenticidade do documento de segurança (38); e em que o documento de segurança é considerado autêntico se a mensagem secreta correta (34) é validada e igual à mensagem secreta esperada.

12. O método implementado por computador para a descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo a reivindicação anterior, **caracterizado por** o módulo descodificador (29) ser executado através de pelo menos uma rede especial transformadora e pelo menos uma rede neuronal convolucional.

13. O método implementado por computador para a descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo com qualquer uma das reivindicações 11 e 12, **caracterizado por** a imagem integral capturada a validar (24) ser obtida a partir da captura de uma imagem de segurança incluída no documento de segurança a validar (22) através de um módulo de captura digital de uma imagem integral incluída em pelo menos um documento de segurança a validar (23).

14. O método implementado por computador para a descodificação e validação da integridade de um documento de segurança com uma imagem codificada por esteganografia resistente à impressão, de acordo com qualquer uma das reivindicações 11 a 13, **caracterizado por** a imagem integral capturada a validar (24) ser um retrato facial e o documento de segurança a validar (22) ser selecionado do grupo

consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

15. Um documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21), **caracterizado por** compreender pelo menos uma imagem integral codificada (17) transmitida fisicamente num suporte físico ou num suporte eletrónico e ser preparado por meio do método definido em qualquer uma das reivindicações 8 a 10.

16. O documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21), de acordo com a reivindicação anterior, **caracterizado por** a imagem integral codificada (17) ser um retrato facial e o documento de segurança com uma imagem codificada por esteganografia resistente à impressão (21) ser selecionado do grupo consistindo num documento de identificação pessoal e um documento de viagem pessoal legível por máquina.

17. Um aparelho computacional, **caracterizado por** compreender meios adaptados para executar as etapas do método definido em qualquer uma das reivindicações 8 a 10.

18. Um aparelho computacional, **caracterizado por** compreender meios adaptados para executar as etapas do método definido em qualquer uma das reivindicações 11 a 14.

19. Um programa de computador, **caracterizado por** compreender instruções para propiciar que o aparelho computacional, conforme definido na reivindicação 17, execute as etapas do método definido em qualquer uma das reivindicações 8 a 10.

20. Um programa de computador, **caracterizado por** compreender instruções para propiciar que o aparelho computacional, conforme definido na reivindicação 18, execute as etapas do método definido em qualquer uma das reivindicações 11 a 14.

21. Um meio de leitura por um aparelho computacional, **caracterizado por** compreender a instalação do programa de computador, conforme definido na reivindicação 19.

22. Um meio de leitura por um aparelho computacional, **caracterizado por** compreender a instalação do programa de computador, conforme definido na reivindicação 20.

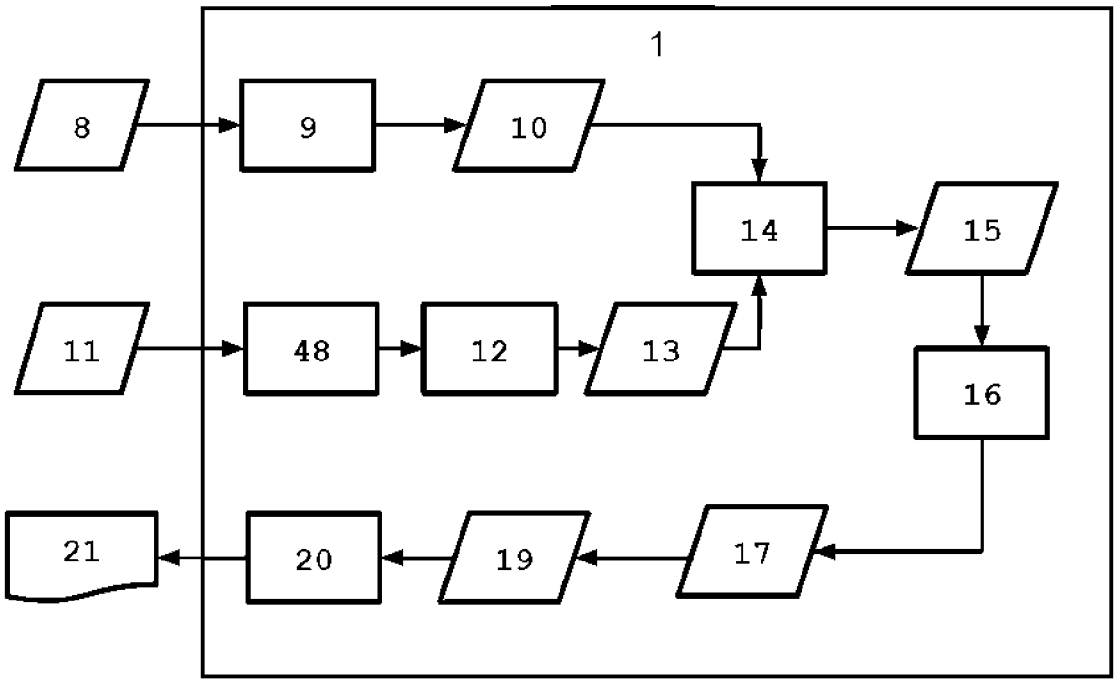

1
8
9
10
14
15
11
48
12
13
16
21
20
19
17

Fig. 1

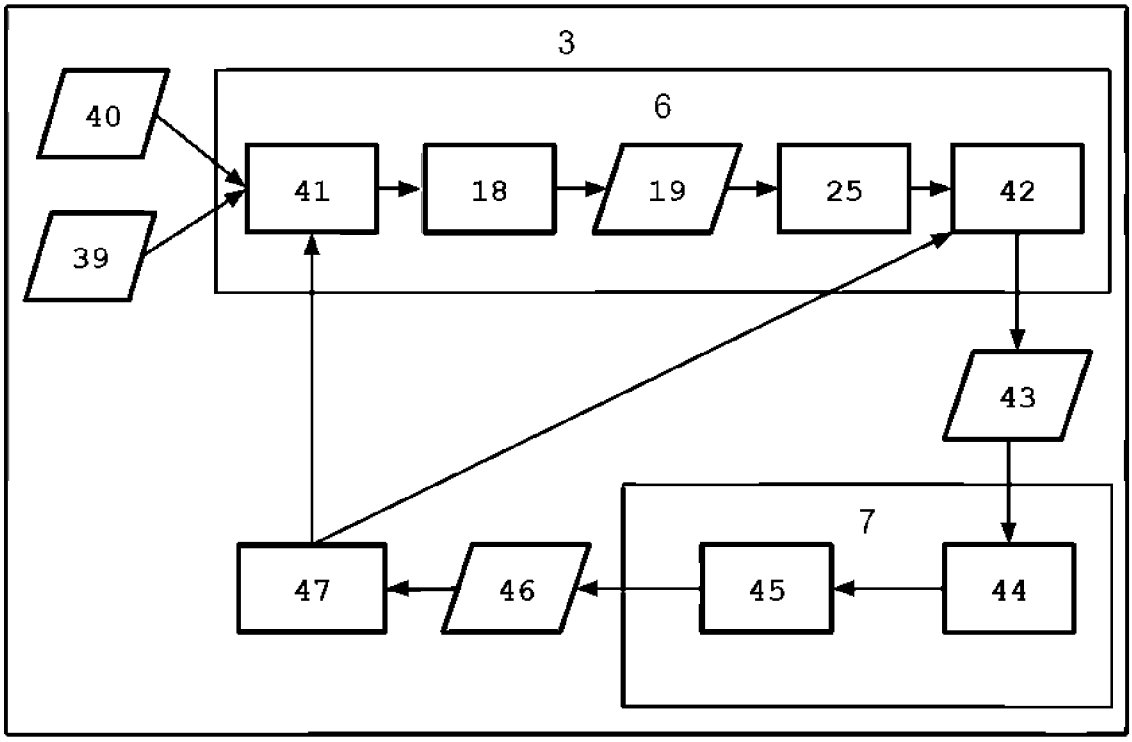

3
6
40
41
18
19
25
42
39
43
7
47
46
45
44

Fig. 2

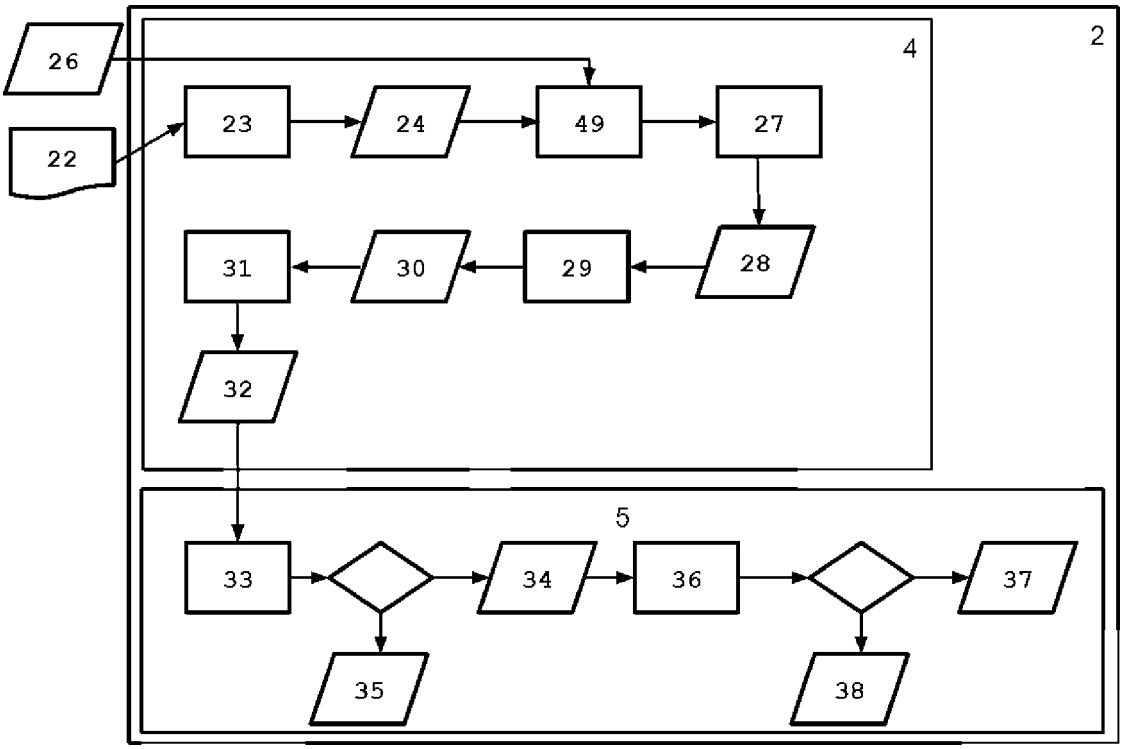


Fig. 3