

confidence score, and comparing the confidence score with a threshold to determine whether the ID card is tampered. Training the neural network comprises inputting each training sample of a plurality of training samples into a binary classifier for a classification based on a face photo of said training sample. The training samples include a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo, and a plurality of negative training samples obtained from composite images, each composite image being generated by overlaying a face photo on an ID card photo based on an automatic image generation process. Training the neural network further comprises iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples, inputting each composite image of the plurality of composite images into the binary classifier to identify hard negative samples, the hard negative samples being composite images of the plurality of composite images classified by the binary classifier as positive training samples, and iteratively updating the parameters of the neural network based on the hard negative samples to retrain the binary classifier.

[0006] Another embodiment provides a system for tamper detection comprising at least one processor, and a computer-readable memory coupled to the processor and having instructions stored thereon. The instructions are executable by the processor to extract image data of a face area of an identity (ID) card from an image of the ID card, apply a trained neural network to the extracted image data to calculate a confidence score, and determine whether the ID card is tampered based on a comparison of the confidence score with a threshold. The trained the neural network comprises a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and neural network parameters iteratively optimized based on classification results of the plurality of training samples. The training samples include a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo, and a plurality of negative training samples obtained from composite images, each composite image comprising a face photo overlaid on an ID card photo based on an automatic image generation. Each composite image of the plurality of composite images is inputted into the binary classifier to identify hard negative samples, the hard negative samples being composite images of the plurality of composite images classified by the binary classifier as positive training samples. The parameters of the neural network are iteratively updated based on the hard negative samples to retrain the binary classifier.

[0007] Another embodiment provides an apparatus comprising a receiving device for receiving an image of an identity (ID) card, and a processing device for extracting image data of a face area of the ID card from the received image. The processing device provides the extracted image data to a trained neural network to calculate a confidence score and determines whether the ID card is tampered based on a comparison of the confidence score with a threshold. The trained neural network comprises a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and neural network parameters iteratively optimized based on classification results of the plurality of training samples. The training samples include a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo, and a plurality of negative training samples obtained from composite images, each composite image comprising a face photo overlaid on an ID card photo based on an automatic image generation. Each composite image of the plurality of composite images is inputted into the binary classifier to identify hard negative samples, the hard negative samples being composite images of the plurality of composite images classified by the binary classifier as positive training samples. The parameters of the neural network are iteratively updated based on the hard negative samples to retrain the binary classifier.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Embodiments of the disclosure will be better understood and readily apparent to one of ordinary skill in the art from the following written description, by way of example only, and in conjunction with the drawings, in which:

CLAIMS

1. A tamper detection method comprising:
 - extracting image data of a face area of an identity (ID) card from an image of the ID card;
 - applying a trained neural network to the extracted image data to calculate a confidence score, wherein training the neural network comprises:
 - inputting each training sample of a plurality of training samples into a binary classifier for a classification based on a face photo of said training sample, wherein the training samples comprise:
 - a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo; and
 - a plurality of negative training samples obtained from composite images, wherein each composite image is generated by overlaying a face photo on an ID card photo based on an automatic image generation process;
 - iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples;
 - inputting each composite image of the plurality of composite images into the binary classifier to identify hard negative samples, wherein the hard negative samples are composite images of the plurality of composite images classified by the binary classifier as positive training samples; and
 - iteratively updating the parameters of the neural network based on the hard negative samples to retrain the binary classifier; and
 - comparing the confidence score with a threshold to determine whether the ID card is tampered.
2. The method as claimed in claim 1, wherein the confidence score is calculated using a softmax function.
3. The method as claimed in claim 1 or 2, wherein the threshold is determined from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.01.

4. The method as claimed in claim 1 or 2, wherein the threshold is determined from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.001.
5. The method as claimed in claim 1, wherein overlaying the face photo on the ID card photo based on the automatic image generation process comprises:
 - randomly selecting an image having the face photo and an image having the ID card photo from respective image libraries;
 - extracting the face photo and the ID card photo from the selected images, wherein the ID card photo comprises a region of interest where a face photo is located;
 - adjusting a size and a brightness of the extracted face photo based on the size and the brightness of the region of interest of the ID card photo; and
 - overlaying the size- and brightness-adjusted face photo onto the region of interest of the ID card photo.
6. The method as claimed in claim 5, adjusting the size and the brightness of the extracted face photo further comprises adding shadow at a random angle to the extracted face photo.
7. The method as claimed in claim 5 or 6, further comprising adding Gaussian noise and Gaussian blur to the composite image.
8. The method as claimed in any one of the preceding claims, wherein each training sample comprises a rectangular shape, and wherein the face area of the ID card comprises a rectangular shape of the same size as the training sample.
9. A system for tamper detection, comprising:
 - at least one processor; and
 - a computer-readable memory coupled to the processor and having instructions stored thereon that are executable by the processor to:
 - extract image data of a face area of an identity (ID) card from an image of the ID card;
 - apply a trained neural network to the extracted image data to calculate a confidence score, wherein the trained the neural network comprises:

a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, wherein the training samples comprise:

a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo; and

a plurality of negative training samples obtained from composite images, wherein each composite image comprises a face photo overlaid on an ID card photo based on an automatic image generation, and

neural network parameters iteratively optimized based on classification results of the plurality of training samples,

wherein each composite image of the plurality of composite images is inputted into the binary classifier to identify hard negative samples, the hard negative samples being composite images of the plurality of composite images classified by the binary classifier as positive training samples; and

wherein the parameters of the neural network are iteratively updated based on the hard negative samples to retrain the binary classifier; and

determine whether the ID card is tampered based on a comparison of the confidence score with a threshold.

10. The system as claimed in claim 9, wherein the instructions are executable by the processor to calculate the confidence score using a softmax function.

11. The system as claimed in claim 9 or 10, wherein the memory further comprises instructions executable by the processor to determine the threshold from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.01.

12. The system as claimed in claim 9 or 10, wherein the memory further comprises instructions executable by the processor to determine the threshold from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.001.

13. The system as claimed in claim 9, wherein the face photo of each composite image comprises a size and a brightness adjusted based on the size and the brightness of a region of interest of the ID card photo where the face photo is located.

14. The system as claimed in claim 13, wherein the face photo of each composite image comprises an added shadow at a random angle.

15. The system as claimed in claim 13 or 14, wherein each composite image comprises added Gaussian noise and Gaussian blur.

16. The system as claimed in any one of claims 9 to 15, wherein each training sample comprises a rectangular shape, and wherein the face area of the ID card comprises a rectangular shape of the same size as the training sample.

17. An apparatus comprising:

a receiving device for receiving an image of an identity (ID) card; and

a processing device for extracting image data of a face area of the ID card from the received image, the processing device providing the extracted image data to a trained neural network to calculate a confidence score and determining whether the ID card is tampered based on a comparison of the confidence score with a threshold, wherein the trained neural network comprises:

a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, wherein the training samples comprise:

a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo; and

a plurality of negative training samples obtained from composite images, wherein each composite image comprises a face photo overlaid on an ID card photo based on an automatic image generation, and

neural network parameters iteratively optimized based on classification results of the plurality of training samples,

wherein each composite image of the plurality of composite images is inputted into the binary classifier to identify hard negative samples, the

hard negative samples being composite images of the plurality of composite images classified by the binary classifier as positive training samples; and
wherein the parameters of the neural network are iteratively updated based on the hard negative samples to retrain the binary classifier.