Automated Image Evaluation

Technical Field

[0001] This disclosure generally relates to methods and systems for automated image evaluation.

Background

[0002] This background description is provided for the purpose of generally presenting the context of the disclosure. Contents of this background section are neither expressly nor impliedly admitted as prior art against the present disclosure.

[0003] Cameras are commonly used to capture information in the real world and generate a digital representation of visual information. In particular, cameras in end-user computing devices such as smartphones are often used to image faces, documents, and identification cards etc. that comprise valuable information. However, cameras of various computing devices are configured differently and capture images with different exposure settings. In addition, users capture images under varying lighting conditions and with varying abilities to accurately image relevant artefacts of interest in a document or face.

[0004] Several image processing systems may process captured images to derive meaningful information. The information may include details of a face or facial identity, details of documents such as text, figurative content in identification documents etc. The task of extracting such meaningful information is very computationally intensive because of the huge potential for variability in the cameras, imaging conditions and imaging ability of the individual capturing the images. The variability may lead to the need to capture images of an artefact of interest several times to obtain a sufficiently clear image for automated detection or information extraction. Images poorly suited for automated information extraction may lead to increased computing resource expenditure by downstream image processing systems and a suboptimal user experience in environments where low latency information processing is desirable.

Summary

[0005] Some embodiments relate to a computing device comprising one or more processing units to:

receive a first image from an image source, the image comprising an artefact of interest;

process the first image to generate one or more image evaluation metrics by executing a respective one or more condensed image processing models;

process the one or more image evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

when the degree of clarity is below a predefined threshold, generate a feedback message based on the one or more image evaluation metrics; and

cause presentation of the feedback message on a display in communication with the one or more processing units.

[0006] In some embodiments, the one or more image evaluation metrics are approximations of corresponding image quality metrics.

[0007] In some embodiments, the image quality metrics comprise any one of: blur score, dark/light score, BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) score, NIQE (Naturalness Image Quality Evaluator) score, pixel width score or total variance score.

[0008] In some embodiments, the artefact of interest comprises any one of: a face, an identity card, a document or a photograph.

[0009] In some embodiments, the one or more condensed image processing models generate the one or more image evaluation metrics in real-time or near real-time.

[0010] In some embodiments, the one or more condensed image processing models generate the one or more image evaluation metrics without invoking compute resources remote to the computing device.

[0011] In some embodiments, the computing device is an end user computing device.

[0012] In some embodiments, the one or more processing units further:

receive a second image with the artefact of interest;

perform object detection or text recognition operations on the second image.

[0013] In some embodiments, the one or more processing units further:

identify a region within the first image contributing to the degree of clarity being below the predefined threshold;

cause presentation of the first image visually indicating the identified region on the display.

[0014] Some embodiments relate to a computer-implemented method for image evaluation, comprising:

receiving a first image from an image source, the image comprising an artefact of interest;

processing the first image to generate one or more image evaluation metrics by executing respective one or more condensed image processing models;

processing the one or more image evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

when the degree of clarity is below a predefined threshold, generate a feedback message based on the one or more image evaluation metrics; and

cause presentation of the feedback message on a display.

[0015] Some embodiments relate to a server system comprising one or more processing units to:

transmit to a computing device program code executable by the computing device to perform the method of any one of claims 10 to 18.

[0016] Some embodiments relate to one or more non-transitory computer-readable storage media storing instructions that when executed by one or more processing units to:

receive a first image from an image source, the image comprising an artefact of interest;

process the first image to generate one or more image evaluation metrics by executing respective one or more condensed image processing models;

process the one or more evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

when the degree of clarity is below a predefined threshold, generate a feedback message based on a one or more of image evaluation metrics; and

cause presentation of the feedback message on a display in communication with the one or more processing units.

Brief Description of the Drawings

[0017] Some embodiments of systems and methods for automated image evaluation, in accordance with present disclosure, will now be described, by way of non-limiting example only, with reference to the accompanying drawings in which:

[0018] Figure 1 illustrates a schematic diagram of a method of image evaluation;

[0019] Figures 2A and 2B illustrate images with artefacts of interest;

[0020] Figure 3 illustrates an image with artefacts of interest and results of image evaluation;

[0021] Figure 4 illustrates another image with artefacts of interest and results of image evaluation; and

[0022] Figure 5 illustrates a graph with relationships between a blur score obtained using a condensed blur model and a corresponding uncondensed blur model;

[0023] Figure 6 illustrates images with approximate BRISQUE and pixel width scores;

[0024] Figure 7 illustrates images with approximate histograms of over/under exposure detection results;

[0025] Figure 8 is a block diagram of a system for image evaluation; and

[0026] Figure 9 is a flow diagram of a method for image evaluation.

Detailed Description

[0027] The disclosed devices and methods perform image evaluation operations to generate feedback for users. The feedback is responsive to image quality issues in captured images. The image quality issues are identified based on the image evaluation operations. The disclosed methods are intended to improve the process of recognition of artefacts of interest in images by evaluating images and providing pointed and actionable

feedback in response to any quality issues identified as a result of image evaluation. The embodiments do so in real-time or near real-time with respect to the point of time of capture of images. Some embodiments perform the image evaluation operations on an end-user computing device without the need to access compute resources outside of the computing device to provide improved latency and real-time or near real-time performance.

[0028] Artefacts of interest in images comprise documents, identification cards, faces in photographs etc. The artefacts of interest may be intended for identity verification, document verification or for providing evidence as part of a process such as registration with a service provider network etc. To meet the clarity and quality related needs of an underlying process, it is necessary to adequately capture the artefact of interest such that the necessary information is legible or observable by automated information processing systems. For example, a face in a photograph for identity verification may need to be aligned towards the camera with sufficient visibility of facial features to accurately capture the facial identity of a person. Photographs with faces pointing in an unsuitable angle (for example: photograph 410 of Fig. 4) may not be suitable for identity verification or facial feature extraction. Similarly, identity cards when captured in an unfocused image (examples: photograph 210 of Fig. 2A, photograph 310 of Figure 3) may not be suitable for automated processing when compared to clearly captured images (for example: photograph 220 of Figure 2B).

[0029] An end-user computing device referred to in this disclosure comprises a smartphone device, a tablet device, a laptop device etc. that is used by an end user to capture and process images. Computing device 800 of Fig. 8 illustrates a schematic diagram of one such device. The device 800 comprises one or more processing units 810 with access to one or more condensed image processing models 822, recommendation text generation model 824 and information extraction model 824. The device 800 may comprise camera 840 or may otherwise have access to images captured by a camera 840. The image capture device may be integral with, or unitary with, the end-user computing device, or may be a separate device. Device 800 may be in communication with one or more server system 830 that serve as a back end system for an application executing on the device 800. For example, in embodiments where device 800 is a smartphone, the server system 830 may be a backend application server of a relevant application for image evaluation executing on the device 800. The server system 830 may transmit code or information to the device 800 and may receive information from device

800 obtained after processing images captured by camera 840. 812 relates to a real world object such as a document, identity card or a face intended to be imaged by camera 840. Device 800 also comprises a display 814 that allows presentation of recommendations to a user. Device 800 executes the method of the flowchart of Fig. 9.

Image Evaluation Metrics

[0030] The embodiments receive a first image at step 910 of the method of Fig. 1 and generate image evaluation metrics based on the received image. The image evaluation metrics are approximations of image quality metrics. The image evaluation metrics are generated by condensed image processing models. The image quality metrics may include no-reference image quality metrics including one or more of: blur score, dark/light score (through histogram), BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator), NIQE (Naturalness Image Quality Evaluator), Pixel Width (based on Canny edge detection) and total variance. Figure 6 illustrates some example images and associated approximate BRISQUE and pixel width scores.

[0031] The one or more image evaluation metrics allow the determination of a degree of clarity of an artifact of interest in the first image at step 914. The measure of the degree of clarity relates to the measure of how suitable the image is for automatic information extraction. Values of the image evaluation metrics allow the identification of a set of specific problems in the first image. For example, in images of faces for identity verification, issues such as the image being too dark, too blurry, head being rotated such that facial features are occluded may arise. One image evaluation metric may be a face angle metric that indicates an angle of a face in an image with respect to the camera's field of view. In image 410, the face angle metric indicates that the angle is not suitable for identity verification and a feedback message may be generated at step 916 to prompt the user to correct the angle of the face. Similarly, a blur score or a darkness score may be computed to account for blurriness or darkness of the first image and a relevant feedback may be generated to direct a correction of the clarity impeding condition.

[0032] For tasks such as document or identity card scanning that require finer details for performing character recognition, relevant image evaluation metrics may include BRISQUE metric (Blind/Referenceless Image Spatial Quality Evaluator) and/or pixel width metrics. Based on the computed image evaluation metrics, feedback such as: 'text segments out of focus', 'card held too close', 'card held too far', 'camera may have a plastic protective covering, or the lens need to be wiped' may be generated. In some

embodiments, the image evaluation metrics may relate to one or more specific regions of the image contributing to the reduced degree of clarity. Such embodiments may comprise the step of identifying regions in the first image contributing to the degree of unclarity at step 922. The identified regions may be presented on the display 814 at step 924 to allow a user to appreciate the source of the unclarity and take corrective action when capturing a second image. Figure 3 illustrates one such example where image 310 of an identity card is not sufficiently clear for automated processing. Regions 312, 314 and 316 have been identified as sources of unclarity because of unsuitable imagining conditions. A recommendation text 322 has also been presented referring to the relevant regions. The more specific identification of regions in the first image and the generated recommendation text provides the end user with more intuitive feedback so that the user can take corrective action when capturing the second image.

Condensed Image Processing Models

[0033] The computing device 800 is often subject to compute power and latency constraints. However, several of the image quality metrics require significant compute power and memory for accurate computation. In addition, because the embodiments are intended to capture a wide variety of artefacts of interest, each category of artefact of interest may require its own distinctive image quality metrics for evaluation. The wide variety of imaging conditions and image capture approaches may result in a lot of different types of quality issues that may require significant compute resources to accurately and comprehensively detect. The embodiments address these challenges in image evaluation by providing condensed image processing models 822 that are proxies or approximations for the full scale of computations required to generate the image quality metrics. The metrics generated by the condensed image processing models 822 are referred to as image evaluation metrics. The image evaluation metrics serve as approximations of image quality metrics computed by full scale models for image evaluation.

[0034] The condensed image processing models 822 are less resource intensive, portable and faster in comparison to full scale models for image quality metric generation. They allow provision of a more responsive user interface for display of recommendations in relation to the first image. The condensed image processing models perform all necessary computations to generate the image evaluation metrics on the computing device 800. They do not require access to external libraries or compute resources outside of the computing device 800 to provide low latency results. This also avoids the need for

use of any network bandwidth for transmission of images from the computing device to remote systems or servers for further processing. In some embodiments, the maximum latency for execution of step 912 is less than or equal to 100 ms thus providing real-time or near real-time results with respect to the receipt of the first image by the computing device.

[0035] The condensed image processing models may be lite models such as machine learning models. For example, a condensed image processing model may be a deep neural network with a limited number of parameters – e.g. a condensed image processing model may be a lightweight model that takes the top X (e.g. 10) parameters that have the greatest contribution to assessing the respective image evaluation metric. The limited number of parameters keeps the size of the models small allowing them to be deployed on a variety of resource constrained computing devices. The condensed image processing models emulate the output characteristics of other full scale image quality metric generation models in a smaller footprint. For example, generation of blur score image quality metric requires implementation of a variance of laplacian operator iterating over all pixels of an image. This operation may need to be implemented in javascript for web, using objective C for iOS, and using kotlin for android. Maintaining a separate models for the three different environment creates undesirable burdens of fixing bugs or upgrades separately for each of the three platforms.

[0036] Some embodiments include a condensed image processing model for generating a blur score using a deep learning model with a reduced number of parameters (implemented using keras for example). In some embodiments, the condensed image processing model for generating a blur score required 10k parameters with a 10ms runtime. This enabled real-time or near real-time latency in the user experience in image evaluation. The heat map 510 of Fig. 5 illustrates the relationship between blur scores computed by a condensed image evaluation model (image evaluation metric) and blur scores generated by a full scale model (image quality metric). As is reflected in the heat map 510, the condensed image evaluation model largely linearly emulates the scores generated by a full scale model. Similarly, the heat maps of Figure 7 illustrate the relationship between the condensed and full scale models for over/under exposure (image too dark, too bright) evaluation. As illustrated in the histograms of Fig. 7, the condensed image evaluation models closely emulate the results of full scale models while operating with a smaller parameter and computational foot print at low latency. Further, advantageously, the condensed image evaluation models operate

10-20 times faster than the full scale models and are easier to deploy and maintain over heterogeneous platforms.

Feedback Generation

[0037] Based on the image evaluation metrics computed at 912, a decision engine 130 of Fig. 1 may assess whether the first image is suitably clear for subsequent processing. The decision engine receives the plurality of image evaluation metrics (120) generated by the condensed image processing models. In some embodiments, the decision engine may comprise predefined rules or thresholds to determine a degree of clarity of the artefact of interest in the first image. Alternatively, the decision engine may comprise a machine learning model that identifies whether artefact of interest in the first image are captured with a sufficient degree of clarity.

[0038] When the decision engine determines that the degree of clarity is below a predefined threshold or otherwise generally insufficient for extracting information from the first image, then a feedback text or suggestion is generated at step 916. The feedback text or suggestion is generated based on the plurality of image evaluation metrics. Examples of the feedback text include: 'Card held too close', 'Card is obscured by finger', 'Face needs more lighting', 'Bright light/window in the background.' The feedback text is generated by a feedback text generation model 824 that processes the image evaluation metrics, identifies a cause of the unclarity and generates/selects a suitable text describing the unclarity. The recommendation text may be obtained from a library of predefined texts associated with each kind of irregularity in the first image. In some embodiments, the text generation model may be implemented as a classification model that classifies an image with unclarity into one or more categories, each category corresponding to a recommendation text associated with an underlying cause of unclarity in the image. The feedback message is presented on the display 814 of the computing device. In response to the feedback, the user may capture a second image of the artefact of interest received at step 926. The second image may be subjected to similar evaluation processes as the first image. If no unclarity is observed in the second image, then it is subjected to further processing at step 928 to extract information. The extracted information may depend on the nature of the artefact of interest in the image. When the artefact of interest is an image of a face, the corresponding facial feature vector presentations may be extracted. When the artefact of interest is a document or identity card, the corresponding text and figures

in the document/identity card may be extracted by one or more information extraction models 826.

[0039] Some embodiments relate to one or more non-transitory computer-readable storage media storing instructions that when executed by one or more processing units to perform the method of image evaluation. Some embodiments relate to a server system (such as server system 830) comprising one or more processing units to transmit to a computing device program code executable by the computing device to perform the method of image evaluation. The server system 830 may comprise a backend server system supporting an application executing on the computing device 800. The server system 830 may transmit or push the program code for image evaluation to the computing device. For example, when the computing device is a smart phone device, the server system may be a backend code repository storing the smart phone application's code. The transmission between the server system and the computing device need not be direct. It may be intermediated by other computer systems that may form part of a computer network 832. For example, deployment of the program code of the image evaluation method may be intermediated by a smart phone application store.

[0040] The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavor to which this specification relates.

[0041] Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

[0042] The scope of this disclosure encompasses all changes, substitutions, variations, alterations, and modifications to the example embodiments described or illustrated herein that a person having ordinary skill in the art would comprehend. The scope of this disclosure is not limited to the example embodiments described or illustrated herein. Moreover, although this disclosure describes and illustrates respective embodiments herein as including particular components, elements, feature, functions, operations, or steps, any of these embodiments may include any combination or

permutation of any of the components, elements, features, functions, operations, or steps described or illustrated anywhere herein that a person having ordinary skill in the art would comprehend. Although this disclosure describes or illustrates particular embodiments as providing particular advantages, particular embodiments may provide none, some, or all of these advantages.

Claims

1. A computing device comprising one or more processing units to:

   receive a first image from an image source, the image comprising an artefact of interest;

   process the first image to generate one or more image evaluation metrics by executing a respective one or more condensed image processing models;

   process the one or more image evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

   when the degree of clarity is below a predefined threshold, identify a region within the first image contributing to the degree of clarity being below the predefined threshold and generate a feedback message based on the one or more image evaluation metrics; and

   cause presentation of the feedback message on a display in communication with the one or more processing units and presentation of the first image visually indicating the identified region on the display.
2. The computing device of claim 1, wherein the one or more image evaluation metrics are approximations of corresponding image quality metrics.
3. The computing device of claim 2, wherein the image quality metrics comprise any one of: blur score, dark/light score, BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) score, NIQE (Naturalness Image Quality Evaluator) score, pixel width score or total variance score.
4. The computing device of any one of claims 1 to 3, wherein the artefact of interest comprises any one of: a face, an identity card, a document or a photograph.
5. The computing device of any one of claims 1 to 4, wherein the one or more condensed image processing models generate the one or more image evaluation metrics in real-time or near real-time.

6. The computing device of any one of claims 1 to 5, wherein the one or more condensed image processing models generate the one or more image evaluation metrics without invoking compute resources remote to the computing device.
7. The computing device any one of claims 1 to 6, wherein the computing device is an end user computing device.
8. The computing device any one of claims 1 to 7 wherein the one or more processing units further:

    receive a second image with the artefact of interest;

    perform object detection or text recognition operations on the second image.
9. A computer-implemented method for image evaluation, comprising:

    receiving a first image from an image source, the image comprising an artefact of interest;

    processing the first image to generate one or more image evaluation metrics by executing respective one or more condensed image processing models;

    processing the one or more image evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

    when the degree of clarity is below a predefined threshold, identifying a region within the first image contributing to the degree of clarity being below the predefined threshold and generating a feedback message based on the one or more image evaluation metrics; and

    causing presentation of the feedback message on a display and presentation of the first image visually indicating the identified region on the display in combination with the feedback message.
10. The method of claim 9, wherein the one or more image evaluation metrics are approximations of corresponding image quality metrics.

11. The method claim 10, wherein the image quality metrics comprise any one of: blur score, dark/light score, BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) score, NIQE (Naturalness Image Quality Evaluator) score, pixel width score and total variance score.
12. The method of any one of claims 9 to 11, wherein the artefact of interest comprises any one of: a face, an identity card, a document or a photograph.
13. The method of any one of claims 9 to 12, wherein the one or more condensed image processing models generate the one or more image evaluation metrics in real-time or near real-time.
14. The method of any one of claims 9 to 13, wherein the one or more condensed image processing models generate the one or more image evaluation metrics without invoking compute resources remote to the computing device.
15. The method of any one of claims 9 to 14, wherein the method is executed by an end user computing device.
16. The method of any one of claims 9 to 15 further comprising:

    receiving a second image with the artefact of interest;

    performing object detection or text recognition operations on the second image.

    ;

    .

17. A server system comprising one or more processing units to:

    transmit to a computing device program code executable by the computing device to perform the method of any one of claims 9 to 16.

18. One or more non-transitory computer-readable storage media storing instructions that when executed by one or more processing units to:

    receive a first image from an image source, the image comprising an artefact of interest;

process the first image to generate one or more image evaluation metrics by executing respective one or more condensed image processing models;

process the one or more evaluation metrics to determine a degree of clarity of the artefact of interest in the first image;

when the degree of clarity is below a predefined threshold, identify a region within the first image contributing to the degree of clarity being below the predefined threshold and generate a feedback message based on a one or more of image evaluation metrics; and

cause presentation of the feedback message on a display in communication with the one or more processing units and presentation of the first image visually indicating the identified region on the display.