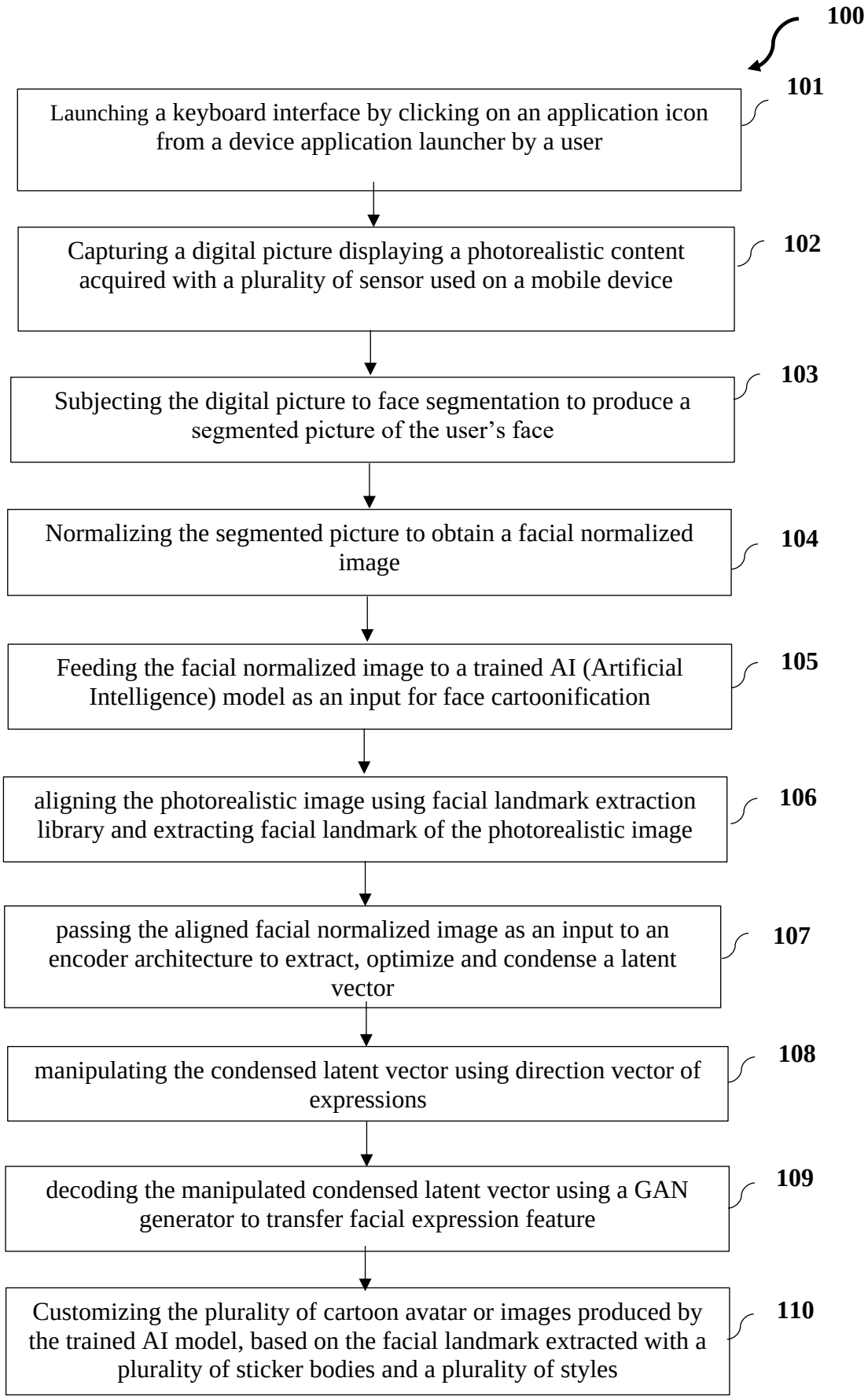

FIG 1

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

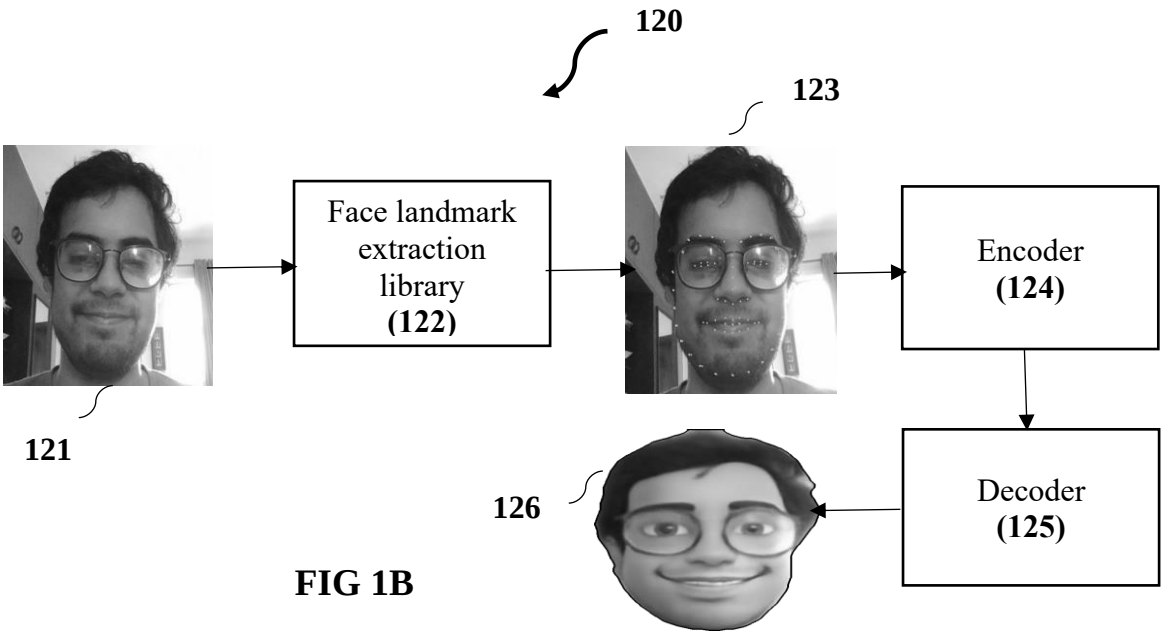

FIG 1B

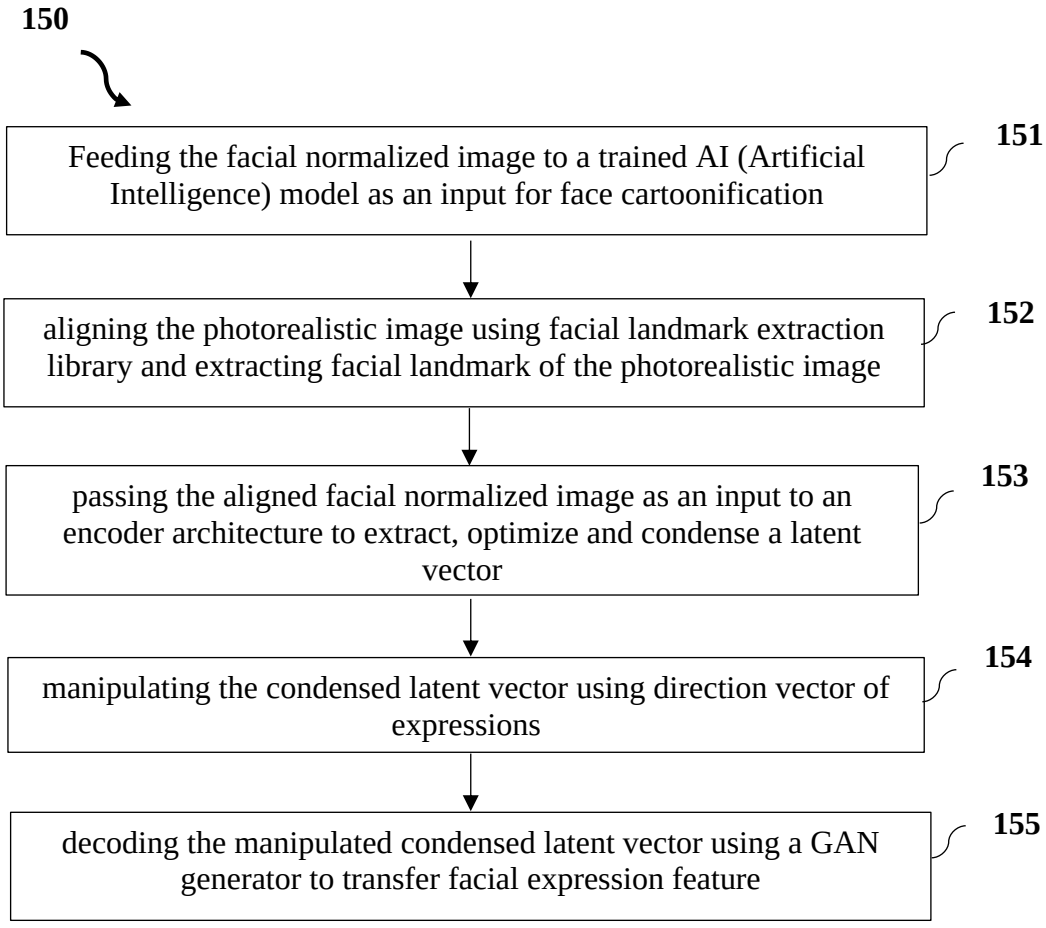

FIG 1C

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

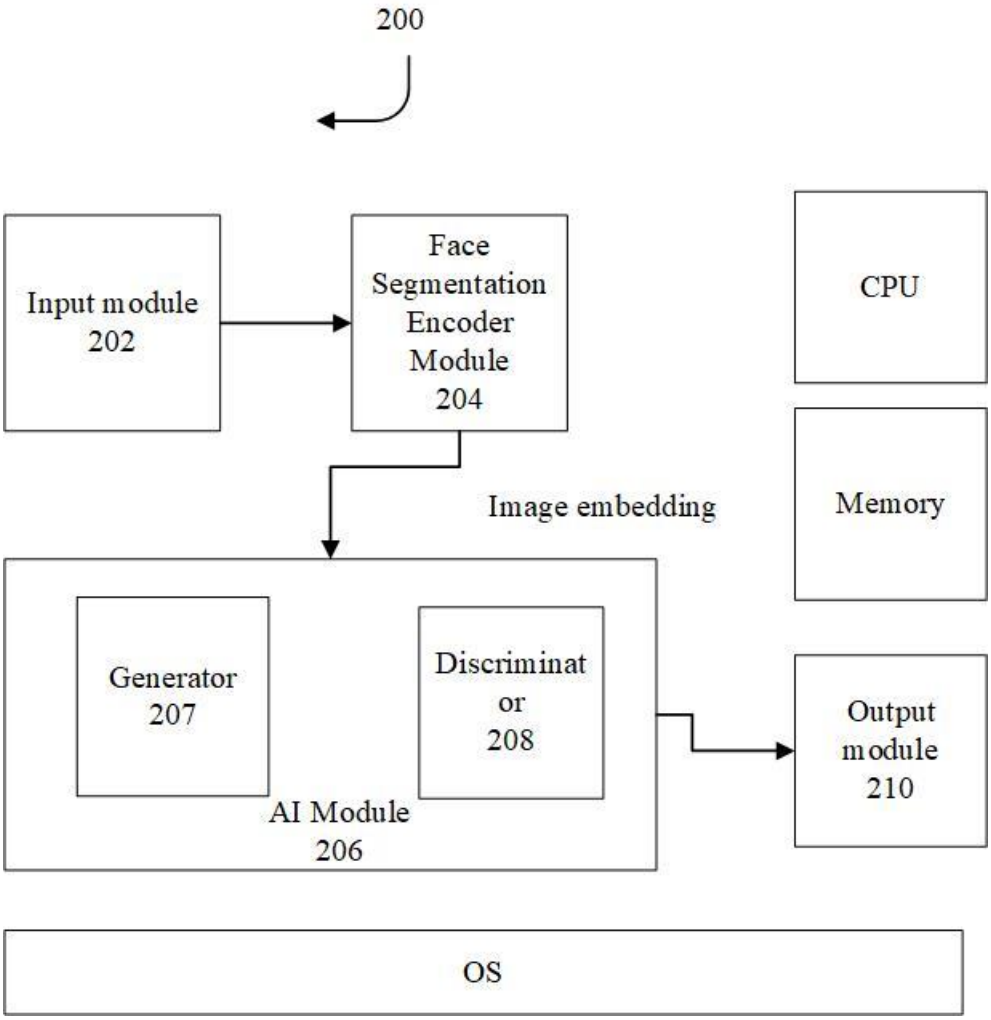

FIG 2

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

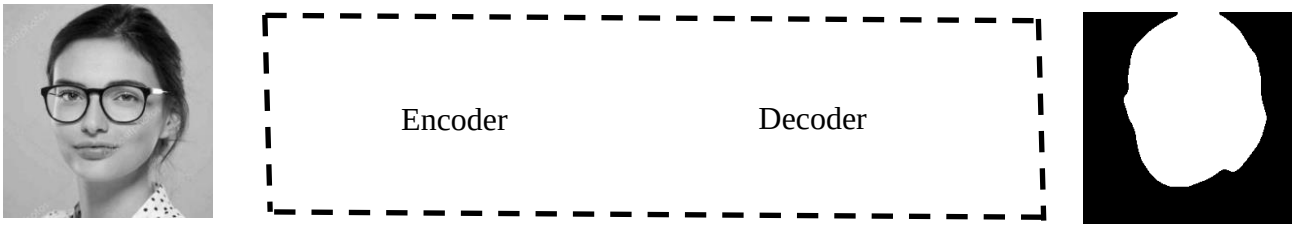

FIG 3

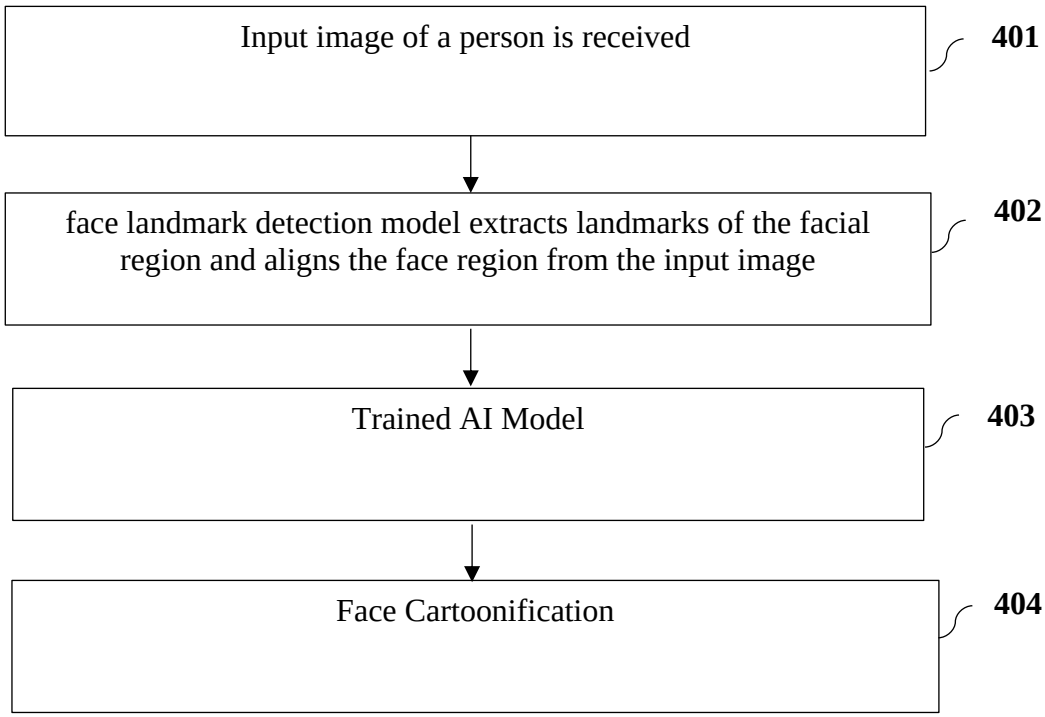

FIG 4

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

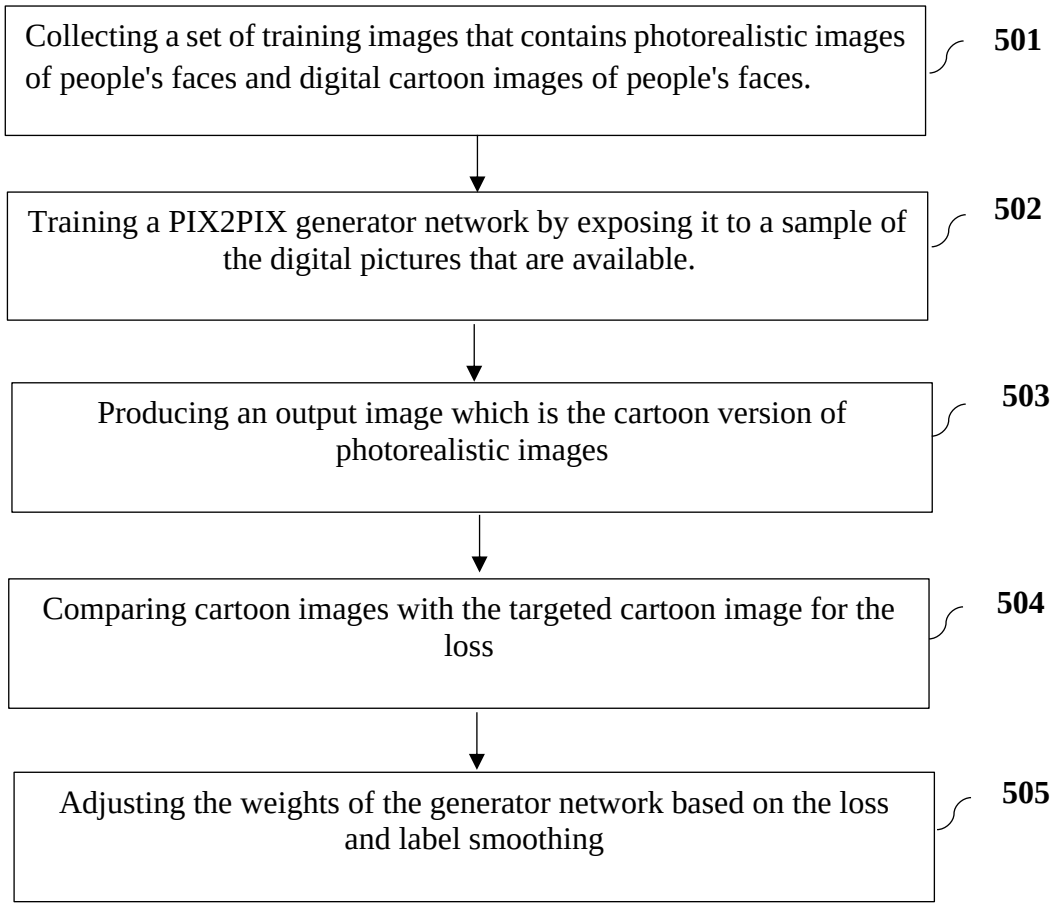

FIG 5

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

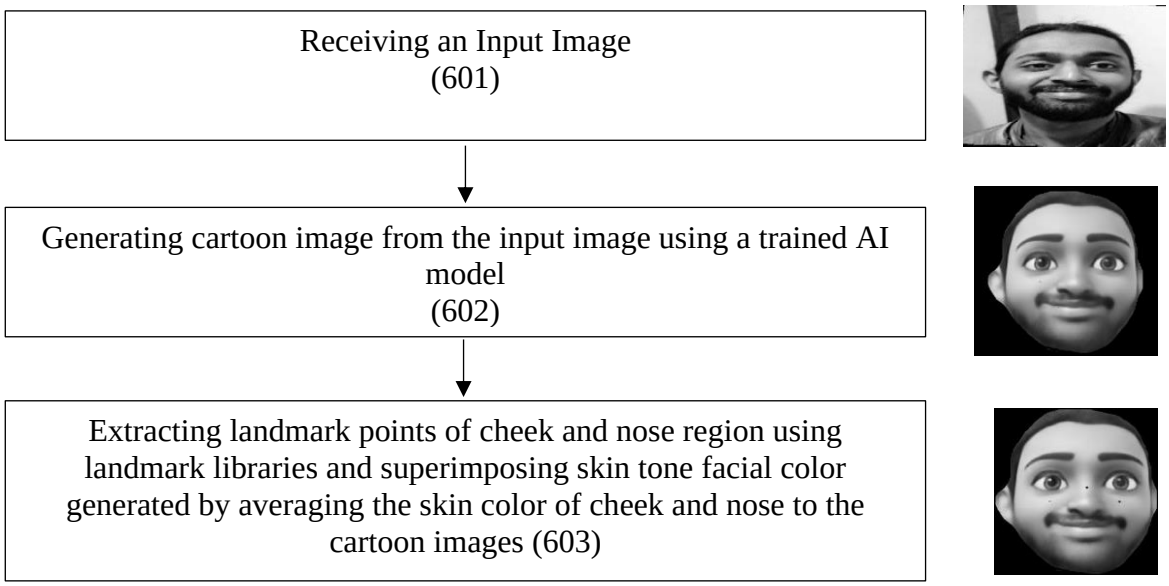

FIG 6



FIG 7

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited



FIG 8A

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

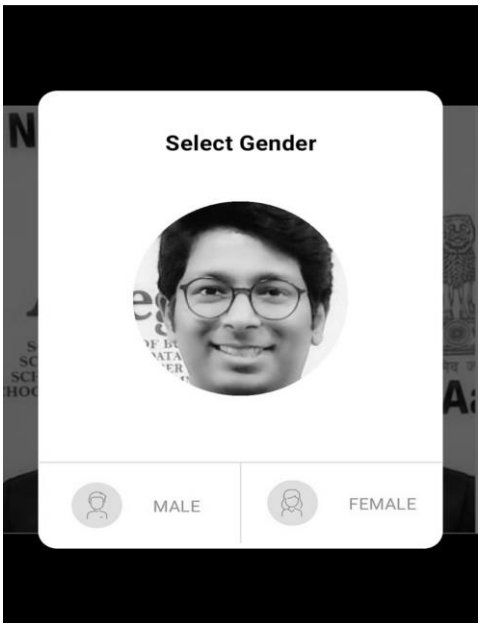

FIG 8B



FIG 8C



FIG 8D

Date: 4th July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

FORM 2

The Patents Act 1970

(39 of 1970)

&

The Patent Rules 2003

COMPLETE SPECIFICATION

(See Section 10 and rule 13)

TITLE

A SYSTEM AND METHOD FOR PERSONALIZED CARTOON IMAGE GENERATION

APPLICANT:

TALENT UNLIMITED ONLINE SERVICES PRIVATE LIMITED

202, S/F 94 Meghdoot Nehru Place,

South Delhi, Delhi, Delhi, India

110019

PREAMBLE:

THE FOLLOWING SPECIFICATION PARTICULARLY DESCRIBES THE INVENTION AND THE MANNER IN WHICH IT IS TO BE PERFORMED

A SYSTEM AND METHOD FOR PERSONALIZED CARTOON IMAGE GENERATION

A) TECHNICAL FIELD

[0001] The present invention is generally related to a field of data processing. The present invention is particularly related to training a neural network to generate cartoon images based on user images. The present invention is more particularly related to a system and method for personalized cartoon image generation, with customized stickers with various expressions.

B) BACKGROUND OF THE INVENTION

[0002] Image is the visual or electronic representation of the person in the social media platform like Facebook, WhatsApp, Instagram. Image is the graphical representation of the subject's facial characteristics. Many digital platforms convert face spatial representation to a two-dimensional form image which is formed by using a few attributes of the subject's face like hair style, color, skin tone, eye color etc., to digitally craft a custom avatar or image similar to subjects. For instance, if the subject image has curly hair, fair skin tone, and blue eyes then an avatar or image is designed based on subject attributes from a plurality of image attribute templates. Like in this case from a pool of hair template curly resemblance will be selected, blue eyes attributes will be selected eye template and face is designed by getting skin tone color from color templates. Furthermore, designed avatars or images are subjected to further deformation based on measured facial attributes and other attributes.

[0003] In addition, avatar or image designed by prior methodology gives identical resemblance for most of the subjects of similar facial attributes or belongs to the same ethnic groups. Furthermore, avatar or image should be more personalized and easily distinguishable. People connected to each other must feel a connection at an emotional level while communicating or sharing publicly, on digital platforms.

[0004] Hence, there is a long-felt need for a system and a method for personalized cartoon image generation, to create an avatar or image that includes most of the facial attributes of the subject which gives resemblance to the actual face of the subject. Still there is a need for an artificial intelligence powered cartoon image generation model are used to generate a personalized avatar or image by ensuring a usage of most of the facial attributes.

[0005] The above-mentioned shortcomings, disadvantages and problems are addressed herein, and which will be understood by reading and studying the following specification.

C) OBJECT OF THE INVENTION

[0006] An object of the present invention is to provide a system and a method for personalized cartoon image generation.

[0007] Another object of the present invention is to provide a system and method for personalized cartoon image generation, to generate an avatar or image that includes most of the facial attributes of the subject which gives resemblance to the actual face of the subject.

[0008] Yet another object of the present invention is to provide a system and method for personalized cartoon image generation, powered by artificial intelligence which makes sure that most of the facial attributes are used to generate a personalized image.

[0009] Yet another object of the present invention is to customize the generated plurality of cartoon image with plurality of styles and plurality of sticker bodies such as storied, gif's and stickers.

[0010] These and other objects and advantages of the present invention will become readily apparent from the following detailed description taken in conjunction with the accompanying drawings.

D) SUMMARY OF THE INVENTION

[0011] The various embodiments of the present invention provide a method and system for personalized cartoon image generation. The method and system disclosed herein involves creation of personalized cartoon image or stickers to make the users part of the conversation and the graphics or content shared look similar to the user face and more aesthetically pleasing instead of using any reference stickers to convey the messages. Furthermore, the method and system utilizes image generation algorithms replicated on the CPU as well as on GPU such that creation of users' personalized heads can be eased and automated.

[0012] According to an embodiment of the present invention, a method for personalized cartoon image generation is provided. The method disclosed herein comprises the steps of launching a bobble keyboard by clicking on an application icon from a device application launcher by a user and capturing a digital picture

displaying a photorealistic content acquired with a plurality of sensors used on a mobile device. The digital picture includes jpeg, png and raw file images. The method further includes subjecting the digital picture to face segmentation to produce a segmented picture of the user's face and normalizing the segmented picture to obtain a facial normalized image. The normalization is performed to scale the segmented picture into a range more familiar or normal to the senses, such that the image looks better for the visualizer. The range of the facial normalized image of pixel intensity values is projected to the predefined range between [-1] to [1]. Furthermore, the method involves feeding the facial normalized image to a trained AI (Artificial Intelligence) model as an input for face cartoonification. The trained AI model learns cartoon facial characteristics that co-relate to a photorealistic image facial feature. In addition, the method includes aligning the facial normalized image fed to the trained AI model as an input using facial landmark extraction library and extracting facial landmark of the photorealistic image by using 68 face landmark points to obtain a plurality of cartoon image. The extraction of facial landmark includes one to one mapping of media-pipe facial points to dlib facial points. Further, the method includes passing the aligned facial normalized image as an input to an encoder architecture to extract, optimize and condense a latent vector of the input. Then, manipulating the condensed latent vector using direction vector of expressions such as smile or sad. Besides, the method includes decoding the manipulated condensed latent vector using a GAN (Generative Adversarial Network) generator to transfer facial expression feature. Finally, the method includes customizing the plurality of cartoon images produced by the trained AI

(Artificial Intelligence) model, based on the facial landmark extracted with a plurality of sticker bodies and a plurality of styles.

[0013] According to an embodiment of the present invention, the face segmentation employs a pretrained mobileunet model which is a combination of mobile net and UNET that has been trained on user photos. The pretrained mobileunet model extracts the face from the backdrop.

[0014] According to an embodiment of the present invention, the facial normalized image as an input to the trained AI model corresponds to a digital image depicting a face of a person. The AI model is trained using the supervised approach, in which a paired data corresponding to photorealistic images and a plurality of cartoon images of the photorealistic images are used for training. Furthermore, the trained AI model is configured as a PIX2PIX GAN to output the plurality of cartoon image of the digital image uploaded by the user.

[0015] According to an embodiment of the present invention, the extraction of facial landmark such as nose, eyes, left cheek, chin and right cheek to identify the face tone in the photorealistic image and to merge the customized sticker body to the cartoon image, and also to add aesthetic features to the cartoon image. The aesthetic features include face glow, smoothening and so on.

[0016] According to an embodiment of the present invention, the customization of the plurality of cartoon image includes extracting and transferring the face tone of the photorealistic image to the generated plurality of cartoon image. The customized plurality of cartoon image with the plurality of sticker bodies includes stories, gifs, and stickers.

[0017] According to an embodiment of the present invention, a system for personalized cartoon image generation is disclosed. The system comprises an input module configured to receive a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device by a user. The digital picture includes jpeg, png and raw file images. A face segmentation encoder module, which is a pre-trained machine learning model configured to receive training image data, to extract face mask from the digital picture uploaded by the user, and to generate condensed feature vectors indicative of the face depicted in the input digital picture. The condensed feature vectors extracted is an image embedding. The training image data includes a pair of photorealistic user images and the cartoon versions of the respective photorealistic user images. The system further includes an AI (Artificial Intelligence) module comprising a generator and a discriminator configured to receive the image embedding as an input and create plurality of cartoon images. The generator and the discriminator in conjunction constitutes an image-to-image translation network. In addition, the system comprises an output module configured to generate a customized plurality of cartoon images comprising a plurality of styles and a plurality of sticker bodies and the plurality of sticker bodies includes stories, gifs, and stickers.

[0018] According to an embodiment of the present invention, the image-to-image translation network is a PIX2PIX GAN, based on a generative adversarial network, which is used to learn the mapping between the real image (display picture) and the target image (cartoon image).

[0019] According to an embodiment of the present invention, the AI module consists of a generator and a discriminator, collectively referred to as an image-to-image translation network. The generator is configured to generate plurality of cartoon images based on the input/target pair. The discriminator is configured to determine whether the input/output pair has a close resemblance to the input/target pair, and also the discriminator is used to compare the input/output pair to the input/target pair, to improve the identity, minimize loss and expression preservation while training the network.

[0020] These and other aspects of the embodiments herein will be better appreciated and understood when considered in conjunction with the following description and the accompanying drawings. It should be understood, however, that the following descriptions, while indicating the preferred embodiments and numerous specific details thereof, are given by way of an illustration and not of a limitation. Many changes and modifications may be made within the scope of the embodiments herein without departing from the spirit thereof, and the embodiments herein include all such modifications.

E) BRIEF DESCRIPTION OF THE DRAWINGS

[0021] The other objects, features, and advantages will occur to those skilled in the art from the following description of the preferred embodiment and the accompanying drawings in which:

[0022] FIG. 1 illustrates a method for personalized cartoon image generation, according to an embodiment of the present invention.

[0023] FIG. 1B illustrates a block-diagram of transferring facial expression feature, according to an embodiment of the present invention.

[0024] FIG. 1C illustrates a flowchart on method for transferring facial expression feature, according to an embodiment of the present invention.

[0025] FIG. 2 illustrates a high-level block diagram of a system for personalized cartoon image generation, according to an embodiment of the present invention.

[0026] FIG. 3 exemplarily illustrates the face segmentation carried out using mobile net and UNET model, according to an embodiment of the present invention.

[0027] FIG. 4 illustrates a flowchart of trained AI model with acquired cartoon facial characteristics co-relating to photorealistic image facial features, according to one embodiment of the present invention.

[0028] FIG. 5 illustrates a flowchart of AI model configured as PIX2PIX GAN to output the plurality of cartoon images, according to an embodiment of the present invention.

[0029] FIG. 6 illustrates a flowchart of customization of the generated plurality of cartoon images, according to an embodiment of the present invention.

[0030] FIG. 7, illustrates a picture of cartoon version vs segmented image according to an embodiment of the present invention.

[0031] FIG. 8A-8D illustrates a screenshot of application on method for personalized cartoon image generation, according to an embodiment of the present invention.

[0032] Although the specific features of the present invention are shown in some drawings and not in others. This is done for convenience only as each feature may be combined with any or all of the other features in accordance with the present invention.

F) DETAILED DESCRIPTION OF THE INVENTION

[0033] In the following detailed description, a reference is made to the accompanying drawings that form a part hereof, and in which the specific embodiments that may be practiced is shown by way of illustration. These embodiments are described in sufficient detail to enable those skilled in the art to practice the embodiments and it is to be understood that the logical, mechanical, and other changes may be made without departing from the scope of the embodiments. The following detailed description is therefore not to be taken in a limiting sense.

[0034] The various embodiments of the present invention provide a method and system for personalized cartoon image generation. The method and system disclosed herein involves creation of personalized cartoon image or stickers to make the users part of the conversation and the graphics or content shared look similar to the user face and more aesthetically pleasing instead of using any reference stickers to convey the messages. Furthermore, the method and system utilizes image generation algorithms replicated on the CPU as well as on GPU such that creation of users' personalized heads can be eased and automated.

[0035] The various embodiments of the present invention provide a method for personalized cartoon image generation. The method disclosed herein comprises

the steps of launching a bobble keyboard by clicking on an application icon from a device application launcher by a user and capturing a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device. The digital picture includes jpeg, png and raw file images. The method further includes subjecting the digital picture to face segmentation to produce a segmented picture of the user's face and normalizing the segmented picture to obtain a facial normalized image. The normalization is performed to scale the segmented picture into a range more familiar or normal to the senses, such that the image looks better for the visualizer. The range of the facial normalized image of pixel intensity values is projected to the predefined range between [-1] to [1]. Furthermore, the method involves feeding the facial normalized image to a trained AI (Artificial Intelligence) model as an input for face cartoonification. The trained AI model learns cartoon facial characteristics that co-relate to a photorealistic image facial feature. In addition, the method includes aligning the facial normalized image fed to the trained AI model as an input using facial landmark extraction library and extracting facial landmark of the photorealistic image by using 68 face landmark points to obtain a plurality of cartoon image. The extraction of facial landmark includes one to one mapping of media-pipe facial points to dlib facial points. Further, the method includes passing the aligned facial normalized image as an input to an encoder architecture to extract, optimize and condense a latent vector of the input. Then, manipulating the condensed latent vector using direction vector of expressions such as smile or sad. Besides, the method includes decoding the manipulated condensed latent vector using a GAN (Generative Adversarial

Network) generator to transfer facial expression feature. Finally, the method includes customizing the plurality of cartoon images produced by the trained AI (Artificial Intelligence) model, based on the facial landmark extracted with a plurality of sticker bodies and a plurality of styles.

[0036] According to one embodiment of the present invention, the face segmentation employs a pretrained mobileunet model which is a combination of mobile net and UNET that has been trained on user photos. The pretrained mobileunet model extracts the face from the backdrop.

[0037] According to one embodiment of the present invention, the facial normalized image as an input to the trained AI model corresponds to a digital image depicting a face of a person. The AI model is trained using the supervised approach, in which paired data corresponding to photorealistic images and cartoon images of the photorealistic images are used for training. Furthermore, the trained AI model is configured as a PIX2PIX GAN to output the plurality of cartoon images of the digital image uploaded by the user.

[0038] According to one embodiment of the present invention, the extraction of facial landmark that includes nose, eyes, left cheek, chin and right cheek to identify the face tone in the photorealistic image and to merge the customized sticker body to the cartoon image, and also to add aesthetic features to the cartoon image. The aesthetic features include face glow, smoothening and so on.

[0039] According to one embodiment of the present invention, the customization of the plurality of cartoon images includes extracting and

transferring the face tone of the photorealistic image to the generated plurality of cartoon image. The customized plurality of cartoon images with the plurality of sticker bodies includes stories, gifs, and stickers.

[0040] According to one embodiment of the present invention, a system for personalized cartoon image generation is disclosed. The system comprises an input module configured to receive a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device by a user. The digital picture includes jpeg, png and raw file images. A face segmentation encoder module, which is a pre-trained machine learning model configured to receive training image data, to extract face mask from the digital picture uploaded by the user, and to generate condensed feature vectors indicative of the face depicted in the input digital picture. The condensed feature vectors extracted is an image embedding. The training image data includes a pair of photorealistic user images and the cartoon versions of the respective photorealistic user images. The system further includes an AI (Artificial Intelligence) module comprising a generator and a discriminator configured to receive the image embedding as an input and create plurality of cartoon images. The generator and the discriminator in conjunction constitutes an image-to-image translation network. In addition, the system comprises an output module configured to generate a customized plurality of cartoon images comprising a plurality of styles and a plurality of sticker bodies and the plurality of sticker bodies includes stories, gifs, and stickers.

[0041] According to one embodiment of the present invention, the image-to-image translation network is a PIX2PIX GAN, based on a generative adversarial

network, that is used to learn the mapping between the real image (display picture) and the target image (cartoon image).

[0042] According to one embodiment of the present invention, the AI module consists of a generator and a discriminator, collectively referred to as an image-to-image translation network. The generator is configured to generate cartoon image based on the input/target pair. The discriminator is configured to determine whether the input/output pair has a close resemblance to the input/target pair, and also the discriminator is used to compare the input/output pair to the input/target pair, to improve the identity, minimize loss and expression preservation while training the network.

[0043] FIG. 1 illustrates a method for personalized cartoon image generation, according to an embodiment of the present invention. With regard to FIG. 1, represents an exemplary method 100 disclosed herein employs a method for personalized cartoon image generation. In the method disclosed herein, comprises the steps of launching a keyboard interface such as bobble keyboard by clicking on an application icon from a device application launcher by a user at step 101 and capturing a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device at step 102. The method further includes subjecting the digital picture to face segmentation to produce a segmented picture of the user's face at step 103 and normalizing the segmented picture to obtain a facial normalized image at step 104. The normalization is performed to scale the segmented picture into a range more familiar or normal to the senses, such that the image looks better for the visualizer. Furthermore, the method involves

feeding the facial normalized image to a trained AI (Artificial Intelligence) model as an input for face cartoonification at step 105. The trained AI model learns cartoon facial characteristics that co-relate to a photorealistic image facial feature. In addition, the method includes aligning the facial normalized image fed to the trained AI model as an input using facial landmark extraction library and extracting facial landmark of the photorealistic image by using 68 face landmark points to obtain a plurality of cartoon images at step 106. The extraction of facial landmark includes one to one mapping of media-pipe facial points to dlib facial points. Further, the method includes passing the aligned facial normalized image as an input to an encoder architecture to extract, optimize and condense a latent vector of the input at step 107. Then, manipulating the condensed latent vector using direction vector of expressions such as smile or sad at step 108. Besides, the method includes decoding the manipulated condensed latent vector using a GAN (Generative Adversarial Network) generator to transfer facial expression feature at step 109. Finally, the method includes customizing the plurality of cartoon images or avatar produced by the trained AI (Artificial Intelligence) model, based on the facial landmark extracted with a plurality of sticker bodies and a plurality of styles at step 110.

[0044] FIG. 1B illustrates a block-diagram of transferring facial expression feature, according to an embodiment of the present invention. In an embodiment, the block diagram 120 illustrates a method of transferring facial expression feature to the plurality of cartoon images. The method 120 illustrates an input image 121 fed to the trained AI model using a face landmark extraction library 122, such that

the method 120 aligns the input image and extracts facial landmark of the input image by using 68 face landmark points to obtain a plurality of cartoon images. Further the method includes passing the aligned input image as an input to an encoder 124 architecture to extract, optimize and condense a latent vector of the input image. Then the condensed latent vector of the input image 121 is manipulated using direction vector of expressions such as smile or sad. Besides, the method 120 involves passing the manipulated condensed latent vector to a decoder 125 using a GAN (Generative Adversarial Network) generator to transfer facial expression feature 126. For instance, the block-diagram 120 illustrates for an input image 121 a person with the beard and smiling face, a cartoonized image with out beard and a smiling face is obtained through facial expression feature.

[0045] FIG. 1C illustrates a illustrates a flowchart on method for transferring facial expression feature, according to an embodiment of the present invention. The method 150 includes feeding the facial normalized image to a trained AI (Artificial Intelligence) model as an input for face cartoonification at step 151. The trained AI model learns cartoon facial characteristics that co-relate to a photorealistic image facial feature. In addition, the method includes aligning the facial normalized image fed to the trained AI model as an input using facial landmark extraction library and extracting facial landmark of the photorealistic image by using 68 face landmark points to obtain a plurality of cartoon images at step 152. The extraction of facial landmark includes one to one mapping of media-pipe facial points to dlib facial points. Further, the method includes passing the aligned facial normalized image as an input to an encoder architecture to extract,

optimize and condense a latent vector of the input at step 153. Then, manipulating the condensed latent vector using direction vector of expressions such as smile or sad at step 154. Besides, the method includes decoding the manipulated condensed latent vector using a GAN (Generative Adversarial Network) generator to transfer facial expression feature at step 155.

[0046] FIG. 2 illustrates a high-level block diagram of a system for personalized cartoon image generation, according to an embodiment of the present invention. In an embodiment, the system 200 comprises an input module 202 configured to receive a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device by a user. The digital picture includes jpeg, png and raw file images. A face segmentation encoder module 204, which is a pre-trained machine learning model configured to receive training image data, to extract face mask from the digital picture uploaded by the user, and to generate condensed feature vectors indicative of the face depicted in the input digital picture. The condensed feature vectors extracted is an image embedding. The training image data includes a pair of photorealistic user images and the cartoon versions of the respective photorealistic user images. The system further includes an AI (Artificial Intelligence) module 206 comprising a generator 207 and a discriminator 208 configured to receive the image embedding as an input and create plurality of cartoon images. The generator 207 and the discriminator 208 in conjunction constitutes an image-to-image translation network. In addition, the system comprises an output module 210 configured to generate a customized

plurality of cartoon images comprising a plurality of styles and plurality of sticker bodies and the plurality of sticker bodies includes stories, gifs, and stickers.

[0047] FIG. 3 exemplarily illustrates the face segmentation carried out using mobile net and UNET model, according to an embodiment of the present invention. The face segmentation is carried out mobileunet model, which is a combination of mobile net and UNET that has been trained on user photos. The encoder 302 utilizes pre-trained classification network like VGG/ResNet, where convolution blocks are followed by a max pool down sampling to encode the input image into feature representations at multiple different levels. Furthermore, the decoder 304 semantically project the discriminative features (lower resolution) learned by the encoder 302 onto the pixel space (higher resolution) to get a dense classification. The decoder 304 consists of up sampling and concatenation followed by regular convolution operations.

[0048] FIG. 4 illustrates a flowchart of trained AI model with acquired cartoon facial characteristics co-relating to photorealistic image facial features, according to one embodiment of the present invention. In an embodiment, an input image of a person is received at step 401. At step 402, the face landmark detection model extracts landmarks of the facial region and aligns the face region from the input image. The digital image that contains only the face region is provided as input to a deep-learning model that has been trained using a supervised approach, to identify the cartoon facial features that correspond to facial features depicted in digital images at step 403. Furthermore, the trained deep-learning model is trained using supervised training. Further, the trained machine learning model is trained,

at least in part, to identify cartoon facial features (e.g., that are part of one or more facial feature landmarks) that correspond to facial features depicted in digital images at step 404. In addition, the parameterized weights are received as output from the machine learning model. In accordance with the principles discussed herein, the parameterized weights describe the cartoon features the deep learning model identified as corresponding to the person's facial features based on the digital image. By way of example, the parameterized cartoon head is received as the output of the trained deep learning model.

[0049] FIG. 5 illustrates a flowchart of AI model configured as PIX2PIX GAN to output the cartoon image, according to an embodiment of the present invention. In an embodiment, a set of training images is collected, that contains photorealistic images of people’s faces and digital cartoon images of people’s faces at step 501. A PIX2PIX GAN generator network is trained by exposing it to a sample of the digital pictures that are available at step 502. The output cartoon image of the generated model and the given input, photorealistic image pair of images is the generated pair or fake pair. Further, the received input image (photorealistic image) is passed through a series of convolution and up-sampling layers. Finally, the method produces an output image which is the cartoon version of photorealistic images at step 503. Furthermore, at step 504, the generated cartoon images are compared with the targeted cartoon image for the loss. Finally, the weights of the generator network is adjusted based on the loss and label smoothing is carried out at step 505. The label smoothing is a regularization technique, which is used to train deep neural networks, such that the technique can improve the

generalization of models effectively. Furthermore, the label smoothing regularization technique also reduces the logit gap, which improves the quality of cartoon image generated. A large logit gap makes the model less adaptive.

[0050] FIG. 6 illustrates a flowchart of customization of the generated cartoon image, according to an embodiment of the present invention. In an embodiment, the face tone of the photorealistic image extracted is customized and transferred to a generated cartoon image. At step 601, an input image is received. Further, at step 602, a cartoon image is generated from the input image (photorealistic image) using a trained AI model. Finally, at step 603, landmark points of cheek and nose region is extracted using landmark libraries and skin tone facial color generated by averaging the skin color of cheek and nose to cartoon images is superimposed.

[0051] Furthermore, customization of the plurality of sticker bodies to the generated plurality of cartoon images is carried out by selecting suitable stickers for the generated cartoon head from the pool of stickers based on the attributes. For instance, while creating the head, attributes such as gender, age sticker bodies are selected, and heads are aligned to the sticker bodies with different conditions based upon the requirements of the stickers. The different conditions include face rotation, face color selection, dress for the body, movement of sticker body to produce gifs and accessories addition on face and sticker body.

[0052] FIG. 7, illustrates a picture of cartoon version vs segmented image according to an embodiment of the present invention. FIG. 7 illustrates a picture 700, in which the first part of the picture is the cartoon version and the second part.

Training data to develop AI models to generate the human body consists of pair images data. Each pair of images contains one segmented human face and one cartoon image created by a graphical illustrator who maintains the resemblance of human features in the cartoon image. Segmentation of the human face is performed by face segmentation model using mobileunet algorithm.

[0053] FIG. 8A-8D illustrates a screenshot of application on method for personalized cartoon image generation, according to an embodiment of the present invention. FIG. 8A illustrates a screenshot, wherein initially the application is launched and the system 200 receives an input image from the user either by clicking an image from the camera or the user can choose an image from a gallery of already saved images. FIG. 8B illustrates a screenshot, wherein the system 200 then receives the gender of the user. Further, cartoonification of the face is done according to the gender of the user. For instance, a male cartoon head will contain mustache and a beard to make it as a cartoon image. Furthermore, the FIG. 8C illustrates the screenshot of final output which cartoon image. Finally, FIG. 8D illustrates a screenshot of customized cartoon image generated with sticker bodies.

[0054] The foregoing description of the specific embodiments will so fully reveal the general nature of the embodiments herein that others can, by applying current knowledge, readily modify and/or adapt for various applications such specific embodiments without departing from the generic concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments.

[0055] It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. Therefore, while the embodiments herein have been described in terms of preferred embodiments, those skilled in the art will recognize that the embodiments herein can be practiced with modifications.

G) ADVANTAGES OF THE INVENTION

[0056] The various embodiments of the present invention provide a system and method for personalized cartoon image generation. The main advantage of the system and method for personalized cartoon image generation is that the method helps to generate customized stickers with various expressions. The present invention also produces more realistic expression and feature, that are copied to user's cartoon characters. In terms of efficiency the system is capable of processing images and generate a number of stickers in less than 500 ms. Furthermore, the present invention works best for images captured in low lightning conditions for different ethnicities.

[0057] The foregoing description of the specific embodiments will so fully reveal the general nature of the embodiments herein that others can, by applying current knowledge, readily modify and/or adapt for various applications such as specific embodiments without departing from the generic concept, and, therefore, such adaptations and modifications should and are intended to be comprehended within the meaning and range of equivalents of the disclosed embodiments.

[0058] It is to be understood that the phraseology or terminology employed herein is for the purpose of description and not of limitation. Therefore, while the

embodiments herein have been described in terms of preferred embodiments, those skilled in the art will recognize that the embodiments herein can be practiced with modifications. However, all such modifications are deemed to be within the scope of the claims.

CLAIMS

We claim:

1. A method (100) for personalized cartoon image generation comprising the steps of:
    a. launching a keyboard interface by clicking on an application icon from a device application launcher by a user (101);
    b. capturing a digital picture displaying a photorealistic content acquired with a plurality of sensors used on a mobile device (102);
    c. subjecting the digital picture to face segmentation to produce a segmented picture of the user's face (103);
    d. normalizing the segmented picture to obtain a facial normalized image (104); and wherein the normalization is performed to scale the segmented picture into a range which is familiar or normal to the senses, such that the visual appearance of image is increased for the visualizer;
    e. feeding the facial normalized image to a trained AI (Artificial Intelligence) model as an input for face cartoonification (105); and wherein the trained AI model learns cartoon facial characteristics that co-relate to a photorealistic image facial feature;
    f. aligning the facial normalized image fed to the trained AI model as an input using facial landmark extraction library and extracting facial landmark of the photorealistic image (106) by using 68 face

landmark points to obtain a plurality of cartoon images; and wherein the extraction of facial landmark includes one to one mapping of media-pipe facial points to dlib facial points

g. passing the aligned facial normalized image to an encoder architecture to extract, optimize and condense a latent vector;

h. manipulating the condensed latent vector using direction vector of expressions; and wherein the direction vector of expressions include smile or sad;

i. decoding the manipulated condensed latent vector using a GAN (Generative Adversarial Network) generator to transfer facial expression feature; and

j. customizing the plurality of cartoon images (107) produced by the trained AI (Artificial Intelligence) model, based on the facial landmark extracted; and wherein the plurality of cartoon images generated is customized with a plurality of sticker bodies and a plurality of styles.

2. The method (100) as claimed in Claim 1, wherein the digital picture includes jpeg, png and raw file images.

3. The method (100) as claimed in Claim 1, wherein the face segmentation employs a pretrained mobileunet model which is a combination of mobile net and UNET that has been trained on user photos; and wherein the pretrained mobileunet model extracts the face from the backdrop.

4. The method (100) as claimed in Claim 1, wherein the range of the facial normalized image of pixel intensity values is projected to the predefined range between [-1] to [1].
5. The method (100) as claimed in Claim 1, wherein the facial normalized image as an input to the trained AI model corresponds to a digital image depicting a face of a person, and wherein the AI model is trained using the supervised approach, in which paired data corresponding to photorealistic images and cartoon images of the photorealistic images are used for training.
6. The method (100) as claimed in Claim 1, wherein the trained AI model is configured as a PIX2PIX GAN to output the cartoon image of the digital image uploaded by the user.
7. The method (100) as claimed in Claim 1, wherein the extraction of facial landmark is done to identify the face tone in the photorealistic image and to merge the customized sticker body to the cartoon image and also to add aesthetic features to the cartoon image, and wherein the facial landmark includes nose, eyes, left cheek, chin and right cheek, and wherein the aesthetic features include face glow and smoothening.
8. The method (100) as claimed in Claim 1, wherein the customization of the cartoon image includes extracting and transferring the face tone of the photorealistic image to the generated cartoon image.
9. The method (100) as claimed in Claim 1, wherein the customized cartoon image with the plurality of sticker bodies includes stories, gifs, and stickers.

10. A system (200) for personalized cartoon image generation comprising:

a. An input module (202) configured to receive a digital picture displaying a photorealistic content acquired with a plurality of sensor used on a mobile device by a user; and wherein the digital picture includes jpeg, png and raw file images;

b. a face segmentation encoder module (204) is a pre-trained machine learning model configured to receive training image data, to extract face mask from the digital picture uploaded by the user, and to generate condensed feature vectors indicative of the face depicted in the input digital picture, and wherein the condensed feature vectors extracted is an image embedding;

c. an AI (Artificial Intelligence) module (206) comprising a generator (207) and a discriminator (208) configured to receive the image embedding as an input and create a plurality of cartoon images, and wherein the generator and the discriminator in conjunction constitutes an image-to-image translation network; and

d. an output module (210) configured to generate a customized plurality of cartoon images comprising a plurality of sticker bodies and plurality of styles; and wherein the plurality of sticker bodies includes stories, gifs, and stickers.

11. The system (200) as claimed in Claim 10, wherein the training image data includes a pair of photorealistic user images and the cartoon versions of the respective photorealistic user images.

12. The system (200) as claimed in Claim 10, wherein the image-to-image translation network is a PIX2PIX GAN, based on a generative adversarial network, that is used to learn the mapping between the real image (display picture) and the target image (cartoon image).

13. The system (200) as claimed in Claim 10, wherein the generator is configured to generate plurality of cartoon images based on the input/target pair.

14. The system (200) as claimed in Claim 10, wherein the discriminator is configured to determine whether the input/output pair has a close resemblance to the input/target pair, and also the discriminator is used to compare the input/output pair to the input/target pair, to improve the identity, minimize loss and expression preservation while training the network.

15. The system (200) as claimed in Claim 10, wherein the digital picture includes jpeg, png and raw file images.

16. The system (200) as claimed in Claim 10, wherein the face segmentation employs a pretrained mobileunet model which is a combination of mobile net and UNET that has been trained on user photos; and wherein the pretrained mobileunet model extracts the face from the backdrop.

17. The system (200) as claimed in Claim 10, wherein the range of the facial normalized image of pixel intensity values is projected to the predefined range between [-1] to [1].

18. The system (200) as claimed in Claim 10, wherein the facial normalized image as an input to the trained AI model corresponds to a digital image

depicting a face of a person; and wherein the AI model is trained using the supervised approach, in which paired data corresponding to photorealistic images and cartoon images of the photorealistic images are used for training.

19. The system (200) as claimed in Claim 10, wherein the trained AI model is configured as a PIX2PIX GAN to output the cartoon image of the digital image uploaded by the user.

20. The system (200) as claimed in Claim 10, wherein the extraction of facial landmark is done to identify the face tone in the photorealistic image and to merge the customized sticker body to the cartoon image, and also to add aesthetic features to the cartoon image, and wherein the facial landmark includes nose, eyes, left cheek, chin and right cheek, and wherein the aesthetic features include face glow and smoothening.

21. The system (200) as claimed in Claim 10, wherein the customization of the cartoon image includes extracting and transferring the face tone of the photorealistic image to the generated cartoon image.

22. The system (200) as claimed in Claim 10, wherein the customized plurality of cartoon images with the plurality of sticker bodies includes stories, gifs, and stickers.

Date: 1st July 2022
Place: Bangalore

Pallavi Oraon
Patent Agent (IN/PA-4066)
FormulateIP Technolegal Solutions Private Limited

A SYSTEM AND METHOD FOR PERSONALIZED CARTOON IMAGE GENERATION

ABSTRACT

The present invention provides a system and method for personalized cartoon image generation. The method (100) comprises launching a keyboard interface (101), capturing a digital picture (102), face segmentation using neural network (103), normalization of segmented face (104), face cartoonification (105), which generates bobble head, facial landmark extraction (106), facial expression feature transfer (109) and customization of the generated plurality of cartoon images (110). Hence, the present invention helps in creation of personalized plurality of cartoon images to make the user part of the conversations and the graphics or content shared look similar to the user input face and more aesthetically pleasing instead of using any reference stickers to convey the messages.

[FIG.1]