



(51) International Patent Classification:

G06F 16/24 (2019.01) G10L 15/26 (2006.01)
G06F 16/33 (2019.01)

(21) International Application Number:

PCT/IB2024/053089

(22) International Filing Date:

29 March 2024 (29.03.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

202321024943 31 March 2023 (31.03.2023) IN

(71) Applicant: **JIO PLATFORMS LIMITED** [IN/IN]; Of-
fice-101, Saffron, Nr. Centre Point, Panchwati 5 Rasta, Am-
bawadi, Ahmedabad, Gujarat 380006 (IN).

(72) Inventors: **JOSHI, Prasad Pradip**; Bungalow #34 & 35,
'Pratisaad', Meadow Gate CHS, Lodha Heaven, Palava,
Dombivli, Maharashtra 421204 (IN). **GUPTA, Naman**;
64-A, Panchwati Colony, Airport Road, Bhopal, Madhya

Pradesh 462030 (IN). **CHATTOPADHYAY, Ritam**; Flat
- 222/B6, Kalyani, Nadia, West Bengal 741235 (IN).

(74) Agent: **KHURANA & KHURANA, ADVOCATES &
IP ATTORNEYS**; D 45, UPSIDC, Site IV, Kasna Road,
Greater Noida, National Capital Region, Uttar Pradesh
201310 (IN).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, CV,

(54) Title: MULTIMODAL MULTILINGUAL SEARCH SYSTEM AND METHOD THEREOF

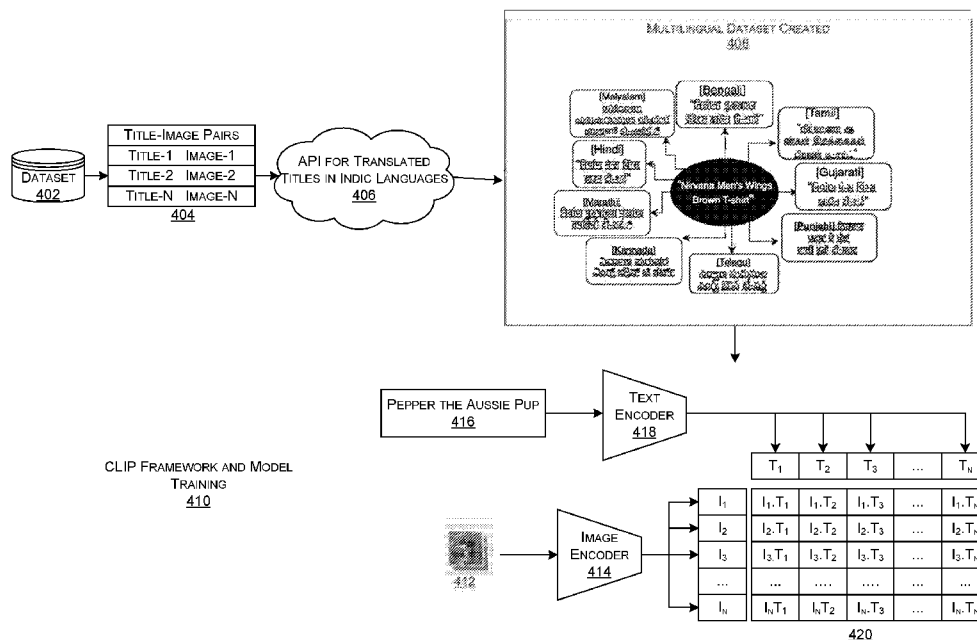


FIG. 4

(57) Abstract: The present disclosure provides a system (110) and a method for enabling multimodal multilingual search in a digital platform. The method includes encoding text data associated with a product in the digital platform in one or more languages, encoding an image associated with the product in the digital platform, and associating the encoded text data with the encoded image to create a multilinguistic training dataset for generating a text-image encoded matrix (408) for the product in the digital platform.

GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- *with international search report (Art. 21(3))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

MULTIMODAL MULTILINGUAL SEARCH SYSTEM AND METHOD THEREOF

RESERVATION OF RIGHTS

[0001] A portion of the disclosure of this patent document contains material, which is
5 subject to intellectual property rights such as, but are not limited to, copyright, design,
trademark, Integrated Circuit (IC) layout design, and/or trade dress protection, belonging to
Jio Platforms Limited (JPL) or its affiliates (hereinafter referred as owner). The owner has no
objection to the facsimile reproduction by anyone of the patent document or the patent
10 disclosure, as it appears in the Patent and Trademark Office patent files or records, but
otherwise reserves all rights whatsoever. All rights to such intellectual property are fully
reserved by the owner.

FIELD OF DISCLOSURE

[0002] The embodiments of the present disclosure generally relate to a search system.
In particular, the present disclosure relates to a multimodal multilingual search system and
15 method thereof.

BACKGROUND OF DISCLOSURE

[0003] The following description of related art is intended to provide background
information pertaining to the field of the disclosure. This section may include certain aspects
of the art that may be related to various features of the present disclosure. However, it should
20 be appreciated that this section be used only to enhance the understanding of the reader with
respect to the present disclosure, and not as admissions of prior art.

[0004] Visual data or image data is getting increased day by day in the internet. A
user enters a text string to search for an image and a search engine needs to find images that
are related to the given text query. Text-to-image retrieval is a challenging task that involves
25 finding images that are related to a given text query.

[0005] Image retrieval is a field of computer vision and natural language processing
that focuses on developing methods for searching and retrieving images from large datasets
based on their visual content and associated information. Contrastive Language-Image Pre-
training (CLIP) and Multilingual Representations for Indian-Languages (MURIL)
30 architectures are related to the field of computer vision and natural language processing
respectively. CLIP handles and processes multimodal aspects of information such as text,
speech, and video. MURIL handles and process multiple languages. A need for a

comprehensive system arises to handle both multimodal and multiple language processing for effective text-to-image processing.

[0006] There is, therefore, a need in the art to provide a method and a system that can overcome the shortcomings of the existing prior arts.

5 OBJECTS OF THE PRESENT DISCLOSURE

[0007] Some of the objects of the present disclosure, which at least one embodiment herein satisfies are as listed herein below.

[0008] It is an object of the present disclosure to enhance image recognition by implementing techniques for better understanding of the content of the images and improve
10 the accuracy of the search results.

[0009] It is an object of the present disclosure to improve cross-language search by allowing users to search for images using text in multiple languages and return results that are relevant to the query regardless of the language used.

[0010] It is an object of the present disclosure to allow users to search based on
15 transliteral queries, where the phonetics of a regional language or a native language is typed in English.

[0011] It is an object of the present disclosure to incorporate metadata, such as image captions, tags, and labels, to improve the relevance of search results.

[0012] It is an object of the present disclosure to enable users to create complex
20 queries, such as including Boolean operators, and return results that match multiple criteria.

[0013] It is an object of the present disclosure to provide more visually appealing and interactive interfaces to improve user experience by allowing users to filter and sort results based on different criteria.

SUMMARY

[0014] In an aspect, the present disclosure relates to a system for enabling multimodal
25 and multilingual search in a digital platform. The system includes one or more processors and a memory operatively coupled to the one or more processors, wherein the memory includes processor-executable instructions, which on execution, cause the one or more processors to encode text data associated with a product in the digital platform in one or more languages,
30 encode an image associated with the product in the digital platform, and associate the encoded text data with the encoded image data to create a multilinguistic training dataset for generating a text-image encoded matrix for the product in the digital platform.

[0015] In some embodiments, the encoded text data may include a translated text and a transliterated text associated with one or more regional languages.

[0016] In some embodiments, the multilinguistic training dataset may include the translated text and the transliterated text corresponding to each of the one or more regional languages associated with the product in the digital platform.

[0017] In some embodiments, the one or more processors may be configured to receive a text query associated with the product in the digital platform from a computing device associated with a user, wherein the text query may include at least one of a foreign language query or a regional language query, and transmit one or more details corresponding to the product based on the text-image encoded matrix to the computing device. In some embodiments, the regional language query may include a transliterated query represented in foreign language.

[0018] In another aspect, the present disclosure relates to a method for enabling multimodal and multilingual search in a digital platform. The method includes encoding, by one or more processors, text data associated with a product in the digital platform in one or more languages, encoding, by the one or more processors, an image data associated with the product in the digital platform, and associating, by the one or more processors, the encoded text data with the encoded image data to generate a multilinguistic training dataset for the product in the digital platform.

[0019] In some embodiments, the method may include receiving, by the one or more processors, a text query associated with the product in the digital platform from a computing device associated with a user, wherein the text query may include at least one of a foreign language query or a regional language query, and transmitting, by the one or more processors, to the computing device one or more details related to the product based on the text-image encoded matrix.

[0020] In another aspect, the present disclosure relates to a user equipment (UE) performing a multimodal and multilingual search. The UE includes one or more processors communicatively coupled to a system, wherein the one or more processors are configured to transmit a text query associated with a product in the digital platform to the system, wherein the text query comprises at least one of a foreign language query or a regional language query and receive, from the system, details related to the product based on the transmitted query and a text-image encoded matrix.

[0021] In yet another aspect, the present disclosure relates to a non-transitory computer readable medium comprising a processor with executable instructions, causing the

processor to encode a text data associated with a product in the digital platform in one or more languages, encode an image associated with the product in the digital platform, and associate the encoded text data with the encoded image to create a multilinguistic training dataset for generating a text-image encoded matrix for the product in the digital platform.

BRIEF DESCRIPTION OF DRAWINGS

[0022] The accompanying drawings, which are incorporated herein, and constitute a part of this disclosure, illustrate exemplary embodiments of the disclosed methods and systems in which like reference numerals refer to the same parts throughout the different drawings. Components in the drawings are not necessarily to scale, emphasis instead being placed upon clearly illustrating the principles of the present disclosure. Some drawings may indicate the components using block diagrams and may not represent the internal circuitry of each component. It will be appreciated by those skilled in the art that disclosure of such drawings includes the disclosure of electrical components, electronic components or circuitry commonly used to implement such components.

[0023] FIG. 1 illustrates an exemplary network architecture (100) in which or with which a proposed system may be implemented, in accordance with some embodiments of the present disclosure.

[0024] FIG. 2 illustrates an exemplary block diagram (200) of a system implemented in the network (100), in accordance with some embodiments of the present disclosure.

[0025] FIG. 3 illustrates an exemplary architecture of a training pipeline (300) in which or with which some embodiments of the present disclosure may be implemented.

[0026] FIG. 4 illustrates an exemplary process flow diagram (400) for text and image embedding and matrix generation implemented in the proposed system, in accordance with some embodiments of the present disclosure.

[0027] FIG. 5 illustrates an exemplary dataset (500) generated using multilingual translation and transliteration in different languages, in accordance with some embodiments of the present disclosure.

[0028] FIGs. 6A-6B illustrate an exemplary query table (600-A, 600-B), respectively, comprising a variety of queries in various languages, in accordance with some embodiments of the present disclosure.

[0029] FIG. 7 illustrates an exemplary computer system (700) in which or with which embodiments of the present disclosure may be implemented.

[0030] The foregoing shall be more apparent from the following more detailed description of the disclosure.

DETAILED DESCRIPTION OF DISCLOSURE

[0031] In the following description, for the purposes of explanation, various specific details are set forth in order to provide a thorough understanding of embodiments of the present disclosure. It will be apparent, however, that embodiments of the present disclosure may be practiced without these specific details. Several features described hereafter can each be used independently of one another or with any combination of other features. An individual feature may not address all of the problems discussed above or might address only some of the problems discussed above. Some of the problems discussed above might not be fully addressed by any of the features described herein.

[0032] The ensuing description provides exemplary embodiments only, and is not intended to limit the scope, applicability, or configuration of the disclosure. Rather, the ensuing description of the exemplary embodiments will provide those skilled in the art with an enabling description for implementing an exemplary embodiment. It should be understood that various changes may be made in the function and arrangement of elements without departing from the spirit and scope of the disclosure as set forth.

[0033] Specific details are given in the following description to provide a thorough understanding of the embodiments. However, it will be understood by one of ordinary skill in the art that the embodiments may be practiced without these specific details. For example, circuits, systems, networks, processes, and other components may be shown as components in block diagram form in order not to obscure the embodiments in unnecessary detail. In other instances, well-known circuits, processes, algorithms, structures, and techniques may be shown without unnecessary detail in order to avoid obscuring the embodiments.

[0034] Also, it is noted that individual embodiments may be described as a process which is depicted as a flowchart, a flow diagram, a data flow diagram, a structure diagram, or a block diagram. Although a flowchart may describe the operations as a sequential process, many of the operations can be performed in parallel or concurrently. In addition, the order of the operations may be re-arranged. A process is terminated when its operations are completed but could have additional steps not included in a figure. A process may correspond to a method, a function, a procedure, a subroutine, a subprogram, etc. When a process corresponds to a function, its termination can correspond to a return of the function to the calling function or the main function.

[0035] The word “exemplary” and/or “demonstrative” is used herein to mean serving as an example, instance, or illustration. For the avoidance of doubt, the subject matter disclosed herein is not limited by such examples. In addition, any aspect or design described herein as “exemplary” and/or “demonstrative” is not necessarily to be construed as preferred or advantageous over other aspects or designs, nor is it meant to preclude equivalent exemplary structures and techniques known to those of ordinary skill in the art. Furthermore, to the extent that the terms “includes,” “has,” “contains,” and other similar words are used in either the detailed description or the claims, such terms are intended to be inclusive—in a manner similar to the term “comprising” as an open transition word—without precluding any additional or other elements.

[0036] Reference throughout this specification to “one embodiment” or “an embodiment” or “an instance” or “one instance” means that a particular feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment of the present disclosure. Thus, the appearances of the phrases “in one embodiment” or “in an embodiment” in various places throughout this specification are not necessarily all referring to the same embodiment. Furthermore, the particular features, structures, or characteristics may be combined in any suitable manner in one or more embodiments.

[0037] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting of the disclosure. As used herein, the singular forms “a,” “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprises” and/or “comprising,” when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof. As used herein, the term “and/or” includes any and all combinations of one or more of the associated listed items.

[0038] Certain terms and phrases have been used throughout the disclosure and will have the following meanings in the context of the ongoing disclosure.

[0039] The term “natural language processing” or “NLP” may refer to a field of artificial intelligence (AI) which provides computers ability to process human language in the form of text or voice data and to ‘understand’ its full meaning, complete with the speaker or writer’s intent and sentiment.

[0040] The term “computer vision” may refer to a field of AI that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs and take actions or make recommendations based on that information. Computer vision enables computer to see, observe and understand.

5 **[0041]** The term “real time” may refer to a level of processing responsiveness that a user or system senses as sufficiently immediate for a particular process or determination to be made, or that enables a processor to keep up with some external process.

[0042] The term “multilingual” may refer to many languages.

[0043] The term “multimodal” may refer to the ability to handle and process different modalities such as text, speech, and video.

10 **[0044]** The term “ResNet-50” may refer to a pre-trained Deep Learning model for image classification of the Convolutional Neural Network (CNN, or ConvNet), which is a class of deep neural networks, most commonly applied to analyzing visual imagery. “ResNet-50” is 50 layers deep and is trained on a million images of 1000 categories from the ImageNet database.

[0045] The term “CLIP” may refer to Contrastive Language–Image Pre-training for learning visual concepts from natural language supervision.

[0046] The term “MURIL” may refer to Multilingual Representations for Indian Languages and is an open-source machine learning tool specifically designed for Indian languages.

20 **[0047]** The term “InfoNCE” where NCE stands for Noise-Contrastive Estimation may refer to a type of contrastive loss function used in self-supervised learning.

[0048] The term “BERT” may refer to Bidirectional Encoder Representations from Transformers (BERT) comprising a masked-language model.

25 **[0049]** The term “IndicBERT” may refer to BERT model specific to Indian languages.

[0050] The term “translation” may refer to the act, process, or product of rendering the meaning of a text or communication from one language into another language.

[0051] The term “transliteration” may refer to the process of writing words using a different alphabet based on the phonetic similarity.

30 **[0052]** The various embodiments throughout the disclosure will be explained in more detail with reference to FIGs. 1-7.

[0053] FIG. 1 illustrates an exemplary network architecture (100) in which or with which embodiments of the present disclosure may be implemented.

[0054] Referring to FIG. 1, the network architecture (100) may include one or more computing devices (104-1, 104-2...104-N) associated with one or more users (102-1, 102-2...102-N) deployed in an environment. A person of ordinary skill in the art will understand that one or more users (102-1, 102-2...102-N) may be individually referred to as the user (102) and collectively referred to as the users (102). Further, a person of ordinary skill in the art will understand that one or more computing devices (104-1, 104-2...104-N) may be individually referred to as the computing device (104) and collectively referred to as the computing devices (104).

[0055] In an embodiment, each computing device (104) may interoperate with every other computing device (104) in the network architecture (100). In an embodiment, the computing devices (104) may be referred to as a user equipment (UE). A person of ordinary skill in the art will appreciate that the terms “computing device(s)” and “UE” may be used interchangeably throughout the disclosure.

[0056] In an embodiment, the computing devices (104) may include, but are not limited to, a handheld wireless communication device (e.g., a mobile phone, a smart phone, a phablet device, and so on), a wearable computer device (e.g., a head-mounted display computer device, a head-mounted camera device, a wristwatch computer device, and so on), a Global Positioning System (GPS) device, a laptop computer, a tablet computer, or another type of portable computer, a media playing device, a portable gaming system, and/or any other type of computer device (104) with wireless communication capabilities, and the like. In an embodiment, the computing devices (104) may include, but are not limited to, any electrical, electronic, electro-mechanical, or an equipment, or a combination of one or more of the above devices such as virtual reality (VR) devices, augmented reality (AR) devices, laptop, a general-purpose computer, desktop, personal digital assistant, tablet computer, mainframe computer, or any other computing device, wherein the computing device (104) may include one or more in-built or externally coupled accessories including, but not limited to, a visual aid device such as camera, audio aid, a microphone, a keyboard, and input devices for receiving input from a user (102) such as touch pad, touch enabled screen, electronic pen, and the like.

[0057] A person of ordinary skill in the art will appreciate that the computing devices or UEs (104) may not be restricted to the mentioned devices and various other devices may be used.

[0058] Referring to FIG. 1, the computing devices (104) may communicate with a system (110), for example, a search and retrieval system through a network (106). In an

embodiment, the network (106) may include at least one of a Fourth Generation (4G) network, a Fifth Generation (5G) network, or the like. The network (106) may enable the computing devices (104) to communicate between devices (104) and/or with the system (110). As such, the network (106) may enable the computing devices (104) to communicate with other computing devices (104) via a wired or wireless network. The network (106) may include a wireless card or some other transceiver connection to facilitate this communication. In an exemplary embodiment, the network (106) may incorporate one or more of a plurality of standard or proprietary protocols including, but not limited to, Wi-Fi, ZigBee, or the like. In another embodiment, the network (106) may be implemented as, or include, any of a variety of different communication technologies such as a wide area network (WAN), a local area network (LAN), a wireless network, a mobile network, a Virtual Private Network (VPN), the Internet, the Public Switched Telephone Network (PSTN), or the like.

[0059] Referring to FIG. 1, the system (110) may include an artificial intelligence (AI) engine (108) in which or with which the embodiments of the present disclosure may be implemented. In particular, the system (110), and as such, the AI engine (108) facilitates search and retrieval of images in the network architecture (100) based on a text query or a search string fed in/typed in by the users (102) on a user interface (UI) of the computing devices (104).

[0060] Further, the system (110) may be operatively coupled to a server (112). In an embodiment, the computing devices (104) may be capable of data communication and information sharing with the server (112) through the network (106). In an embodiment, the server (112) may be a centralised server or a cloud-computing system or any device that is network connected.

[0061] In accordance with an embodiment of the present disclosure, the images searched for by the computing devices (104) may be stored in a database (not shown in FIG. 1). In an embodiment, the image may be stored as text embeddings that may be matched with a text based search query from the user (102) over the computing device (104). Further, a pre-trained image data set may be used to generate text embeddings, wherein the pre-trained image data set may include, such as, but not limited to, image data set associated with fashion, grocery, furniture, books, or any image data set associated with an electronic commerce (e-commerce) site. In an exemplary embodiment, the pre-trained image or pre-trained image data set may include fashion-product images-dataset from a website, wherein the website may include, for example, 45000 unique images and 31000 unique title/description. Further, text conversion includes a translated and transliterated context

associated with various regional or native languages corresponding to a particular region or country. For example, if the website is for India, the text conversions may include different Indian languages, for example, but not limited to, Hindi, Bengali, Punjabi, Gujarati, Tamil, Telugu, Kannada, Marathi, and Malayalam.

5 **[0062]** In an example embodiment, a product catalogue associated with any digital platform such as, but not limited to, an e-commerce website may be used to generate text-based embeddings and a MURIL model may be used to encode product titles and description. A translation application program interface (API) along with look-up based dictionaries may be used to convert the product titles and description into different languages. Each conversion
10 may include a text associated with transliterated context and a text associated with translated context. Further, the texts associated with both the translated and transliterated context may be mapped to a product image presented by a stock keeping unit (SKU) of the e-commerce website to generate image embeddings. In some embodiments, the image embeddings i.e., an image associated with its transliterated text and a related translated text may be used as a
15 training dataset for the AI engine (108). The AI engine (108) may use the training dataset for generating a related image output based on either a translated or transliterated text query. The training dataset may include category-based text representations of the image embeddings. In an example embodiment, the translation API converts the product titles and description into multiple (e.g., nine) different languages, wherein each conversion includes text associated
20 with transliterated context and text associated with translated context, counting to a total of eighteen inputs, that may be mapped to a single product image. In some embodiments, the query may include an audio message and the AI engine (108) may enable generating an image output search based on the audio message.

[0063] In an embodiment, the AI engine (108) may assist in pre-processing of the
25 training dataset. The pre-processing of the text content in the training dataset may be performed by a text encoder, such as, without limitations, a BERT encoder or an IndicBERT, or MURIL encoder and the pre-processing of the image in the training dataset may be performed by an image encoder, such as, without limitations, ResNet-50. Based on the pre-processing and the training dataset, the AI engine (108) may retrieve images of products (e.g.,
30 fashion items) that may be related to the search text entered by the user (102) on the computing device (104).

[0064] In some embodiments, the user (102) may input a regional language based search query for a product (for example, the query may be a text string or an audio message) in the UE (104). The UE (104) may transmit the query to the system (110). The AI engine

(108) may retrieve the images of the product related to the search query. The system (110) may transmit a response for the query as an image related to a product queried to the UE (104) wherein the image of the product may be displayed on a user interface of the UE (104).

[0065] Although FIG. 1 shows exemplary components of the network architecture (100), in other embodiments, the network architecture (100) may include fewer components, different components, differently arranged components, or additional functional components than depicted in FIG. 1. Additionally, or alternatively, one or more components of the network architecture (100) may perform functions described as being performed by one or more other components of the network architecture (100).

[0066] FIG. 2 illustrates an exemplary block diagram (200) of a system (110) implemented for search and retrieval of images in the network (100), in accordance with embodiments of the present disclosure.

[0067] For example, the system (110) may include one or more processor(s) (202). The one or more processor(s) (202) may be implemented as one or more microprocessors, microcomputers, microcontrollers, edge or fog microcontrollers, digital signal processors, central processing units, logic circuitries, and/or any devices that process data based on operational instructions. Among other capabilities, the one or more processor(s) (202) may be configured to fetch and execute computer-readable instructions stored in a memory (204) of the system (110). The memory (204) may be configured to store one or more computer-readable instructions or routines in a non-transitory computer readable storage medium, which may be fetched and executed to create or share data packets over a network service. The memory (204) may comprise any non-transitory storage device including, for example, volatile memory such as Random-Access Memory (RAM), or non-volatile memory such as Electrically Erasable Programmable Read-only Memory (EPROM), flash memory, and the like.

[0068] In an embodiment, the system (110) may include an interface(s) (206). The interface(s) (206) may comprise a variety of interfaces, for example, interfaces for data input and output devices, referred to as input/output (I/O) devices, storage devices, and the like. The interface(s) (206) may facilitate communication for the system (110). The interface(s) (206) may also provide a communication pathway for one or more components of the system (110). Examples of such components include, but are not limited to, processing unit/engine(s) (208) and a database (210).

[0069] The processing unit/engine(s) (208) may be implemented as a combination of hardware and programming (for example, programmable instructions) to implement one or

more functionalities of the processing engine(s) (208). In examples described herein, such combinations of hardware and programming may be implemented in several different ways. For example, the programming for the processing engine(s) (208) may be processor-executable instructions stored on a non-transitory machine-readable storage medium and the hardware for the processing engine(s) (208) may comprise a processing resource (for example, one or more processors), to execute such instructions. In the present examples, the machine-readable storage medium may store instructions that, when executed by the processing resource, implement the processing engine(s) (208). In such examples, the system (110) may include the machine-readable storage medium storing the instructions and the processing resource to execute the instructions, or the machine-readable storage medium may be separate but accessible to the system (110) and the processing resource. In other examples, the processing engine(s) (208) may be implemented by electronic circuitry. In an aspect, the database (210) may comprise data that may be either stored or generated as a result of functionalities implemented by any of the components of the processor (202) or the processing engines (208).

[0070] In an embodiment, the processing engine (208) may include engines that receive simple text queries or Boolean text queries from one or more computing devices via a network such as the computing devices (104) via the network (106) (e.g., via the Internet) of FIG. 1, to retrieve a related image from the database (210). In an embodiment, the processing engine (208) may include one or more modules/engines such as, but not limited to, an acquisition engine (212), an AI engine (214), and other engine(s) (216). A person of ordinary skill in the art will understand that the AI engine (214) may be similar in its functionality with the AI engine (108) of FIG. 1, and hence, may not be described in detail again for the sake of brevity.

[0071] Referring to FIG. 2, the database (210) may include a pre-trained dataset comprising image embeddings and their related text embeddings, wherein the text embedding may include multiple languages and include a translated version and transliterated version of text for each one of the multiple languages. In some embodiments, the multiple languages may include regional or native languages associated with a particular country or region. In an embodiment, the database (210) may or may not reside in the system (110). In an embodiment, the system (110) may be operatively coupled with the database (210).

[0072] In an exemplary embodiment, the text embeddings and the image embedding may be stored in a matrix relation in the database (210). The text embedding may include description or a title associated with a particular image. The description and title may be

available in one or more regional or native languages, for example, Indian languages, wherein the description and title may include both a transliterated text and a translated text associated with each one of the one or more regional or native languages. The text query from the user (102) through a computing device (104) is matched for the related image embeddings in the database (210) and the image matching the text query is sent back to the user (102), which may be viewed on the computing device (104).

[0073] By way of example but not limitation, the one or more processor(s) (202) may generate a multilinguistic dataset based on encoding text data associated with a product in a digital platform, encoding an image associated with the product in the digital platform, and associating the encoded text data with the encoded image. In some embodiments, the multilinguistic dataset may be used to generate a text-image encoded matrix corresponding to the product in the digital platform. Referring to FIG. 2, the one or more processors (202) may receive a search text query for a product in the digital platform from the computing device (104) of FIG. 1 and process the same to retrieve the corresponding image data or product details based on the received search text query from the database (210). In an embodiment, the one or more processor(s) (202) of the system (110) may cause the acquisition engine (212) to extract the data parameters from the multilinguistic dataset stored in the database (210) for analysis by the AI engine (214) for providing a response or result for the received search text query, wherein the response may be transmitted back to the requesting computing device (104). In particular, the data parameters may include a training dataset including one or more detail associated with the product, such as but not limited to, images, manufacturer details, shipping options, pricing, availability, etc. of an e-commerce website and the associated text embeddings. In an embodiment, the one or more processor(s) (202) may cause the AI engine (214) to pre-process the set of data parameters in one or more batches. As described with reference to FIG. 1 above, the AI engine (214) may utilise one or more machine learning models to pre-process and analyze the set of data parameters. In an embodiment, results of the analysis may thereafter be transmitted back to the computing device (104), to other devices, to a server providing a web page to a user (102) of the computing device (104), or to other non-device entities.

[0074] A person of ordinary skill in the art will appreciate that the exemplary (200) may be modular and flexible to accommodate any kind of changes in the system (110). In an embodiment, the pre-trained data may get collected meticulously and deposited in a cloud-based data lake to be processed for image retrieval.

[0075] FIG. 3 illustrates an exemplary architecture of a training pipeline (300) in which or with which some embodiments of the present disclosure may be implemented. The training pipeline (300) includes various components and associated processing steps. The training pipeline (300) provides a curated text translation and transliteration for different languages along with images as training dataset. Further, the training pipeline (300) includes a loss function to train both text and image encoders. The trained dataset is stored as category-based embeddings.

[0076] In FIG. 3, components such as, without limitations, a catalogue dump or catalogue collection (302) component, a database comprising text image pairs (312), a text encoder (316), and an image encoder (336) are shown. Further, FIG. 3 also illustrates the different processing involved with the components. By way of example, and not limitations, the data from the catalogue collection (302) is subjected to one or more of the following procedures pre-processing (304), named entity recognition (NER) tagging (306), translation and transliteration service (308), multilingual text and image pairing (312). The pre-processing (304) includes performing one or more function on the catalogue collection (302) such as, stop word removal, lower casing, and dataset class balancing. The pre-processed data is further sent for NER tagging (306) wherein brands, products and colours are tagged. The NER tagged data is further sent for translation or transliteration service (308) where the text query in any other language is translated to English or transliterated to English to create multilingual text image pairs (312), wherein for example, without limitations, the text (310) may include “Puma black back pack” or “Orange round neck Red Tape T-shirt” and may be associated with a corresponding image (314). The text (310) may further be sent through the text encoder (316) and the image (314) may be sent through the image encoder (336). The text encoder (316) may include at least one of BERT, IndicBERT, or MURIL text encoder. The text encoder (316) includes a positional encoder layer (318), a multi-head attention network layer (320), a first residual connection and LayerNorm layer (322), a feedforward network layer (324), a second residual connection and LayerNorm layer (326), stacked up to 12 times and the output from the stacked layers (318-322) are connected to a maxpool layer (328), followed by a feedforward network (330), and a L2 normalization layer (332). The final output from the encoder (316) is subjected to global embeddings (334) with attract/repel and retrofit variations to obtain a text data that may be aligned with the image data. In an example embodiment, the text encoder (316) includes weights initialized by MURIL and pre-trained on catalogue data.

[0077] On the other hand, the image encoder (336) may include a ResNet-50 encoder. The image encoder (336) includes a stacked convolution neural network (CNN) (338), wherein the CNN is stacked up to 50 times. The stacked CNN (338) is connected to a feedforward network (340) and an L2 normalization layer (342). The output from the text encoder (316) and the image encoder (336) are processed to align (344) correct text-image pairs. Mathematically, the text encoder $T: T \rightarrow R^n$ maps query $q \in T$ in the R^n space, while the image encoder $I: I \rightarrow R^m$. Notably, $I = \{i : i \in R^{(224 \times 224 \times 3)}\}$ for a ResNet-50 model. The image encoder (336) (ResNet-50) projects an input image in R^{2048} [or, $I(i) \in R^{2048} : i \in I$] and the text encoder (316) projects input text in R^{768} [or, $T(q) \in R^{768} : q \in T$], therefore another function to project the vectors in a common embedding space is needed and is achieved by applying an affine transformation over the encoded text/image vectors. In other words, two vectors $h: R^{2048} \rightarrow R^{256}$ and $g: R^{768} \rightarrow R^{256}$ are created with a relation $h(I(i))$ and $g(T(q))$ for $q \in T$ and $i \in I$. Once the text and image vectors share the same embedding space, similar image and text vector pairs are “aligned”, while the dissimilar ones are “non-aligned” the. Similarity is given by $s: R^n \times R^n \rightarrow [0, 1]$ such that $s(u,v) = (u^T v) / (\|u\| \|v\|)$.

[0078] A loss function is implemented in the text and image encoding process based on InfoNCE loss function. The loss function optimizes the loss for a positive pair, given N samples out of which $N - 1$ are negative samples. N here, is defined by the batch size. In a batch of N , N images $(i_1, i_2, \dots, i_N)$ and the corresponding title descriptions $(q_1, q_2, \dots, q_N)$ are fed to the image and text encoder respectively. In this setting, predefined positive pairs are defined by $(i_k, q_k) \forall k \in \{1, 2, \dots, N\}$ leaving the rest of the $N^2 - N$ pairs as negatives. The two loss components defined as:

$$\mathcal{L}_{q|i}(i_k, q_k) = -\log \frac{e^{s(g(T(q_k)), h(I(i_k)))}}{\sum_{j=1}^N e^{s(g(T(q_j)), h(I(i_k)))}}$$

$$\mathcal{L}_{i|q}(i_k, q_k) = -\log \frac{e^{s(g(T(q_k)), h(I(i_k)))}}{\sum_{j=1}^N e^{s(g(T(q_k)), h(I(i_j)))}}$$

The total loss for a given batch of size N is thus:

$$\mathcal{L} = \frac{1}{N} \sum_{k=1}^N (\mathcal{L}_{q|i}(i_k, q_k) + \mathcal{L}_{i|q}(i_k, q_k))$$

[0079] The “aligned” text image pairs are used as pre-set data or training dataset. Referring to FIG. 3, the training pipeline architecture includes the multi-head attention network layer (320) in the text encoder (316) that utilizes attention mechanisms to improve the accuracy of retrieval during a text-image retrieval process. In operation, a text query is encoded using a text encoder (316) and the encoded query is used to attend to the images embeddings in the database (210) of FIG. 2, wherein the image embedding is based on encodings generated from image encoder (336). The attended image embeddings rank the images based on their relevance to the query. Further, the proposed loss function takes into account both the text-image similarity and the diversity of the retrieved images, ensuring that the results are not only relevant but also diverse.

[0080] By way of example, without any limitations, training computation may be performed on a single core P100 graphics processing unit (GPU) for about 15 hours of GPU processing time.

[0081] A person of ordinary skill in the art will appreciate that the architecture of the training pipeline (300) may be modular and flexible to accommodate any kind of changes. Although FIG. 3 shows exemplary components of the architecture of the training pipeline (300), in other embodiments, the architecture of the training pipeline (300) may include fewer components, different components, differently arranged components, or additional functional components than depicted in FIG. 3. Additionally, or alternatively, one or more components of the architecture of the training pipeline (300) may perform functions described as being performed by one or more other components of the architecture of the training pipeline (300).

[0082] FIG. 4 illustrates an exemplary process flow diagram (400) for text and image embedding and matrix generation implemented in the proposed system, in accordance with some embodiments of the present disclosure. In FIG. 4, a dataset (402), a table (404) comprising the details of text-image pairing, an application program interface (API) (406) for translating titles in Indic languages, a multilingual dataset (408) for a particular title is shown. The multilingual dataset (408) or multilingual training dataset may be used for training a machine learning model based on a CLIP framework (410) for generating a text-image encoded matrix (420). In some embodiments, the text-image encoded matrix (420) may be used as look-up table for fetching the product image or product details in a digital platform based on a received search text query. Referring to FIG. 4, a text data (416) is encoded using a text encoder (418) and at the same time an image data (412) corresponding to the text data (416) is encoded by an image encoder (414), the encoded text and image data are further used as a training dataset to train the CLIP framework (410) to generate text-image encoded matrix

(420).a. In some embodiments, based on receiving a search text query, the pre-trained AI engine (108) of FIG. 1 may look-up the text-image encoded matrix (420) to provide a response for the received search text query. In accordance with some embodiments of the present disclosure, the dataset (402) may include product catalogue of any e-commerce site, and the table (404) may include different titles in the product catalogue matched to a corresponding image. The API (406) may include a Translator API for generating translated and transliterated titles in various Indian languages to create the multi-linguistic dataset (408). For example, without limitations, the multilingual training dataset (408) may include the details associated with one or more products, for example, title, image, etc., and a translated text and the transliterated text corresponding to each one of the one or more regional languages. In some embodiments the one or more regional languages may include Indian languages such as, Hindi, Bengali, Punjabi, Gujarati, Tamil, Telugu, Kannada, Marathi, and Malayalam. A text query “Nirvana men’s wings brown T-shirts” is shown translated into different language. By way of example, without limitations if user (102) of FIG. 1 types a query in Tamil “நிர்வாண ஆண்களின் இறக்கைகள் பழுப்பு நிற டி-ஷர்ட்” in the computing device (104) of FIG. 1, the query gets associated with a related query in English “Nirvana men’s wings brown T-shirts” in the database and corresponding image of the queried product is retrieved and displayed to the user (102).

[0083] In some embodiments, the present disclosure relates to semantic information retrieval i.e., representing a single entity or product in multiple languages and images synergizes multilingual and multimodal learning capability of any training network thereby enhancing image retrieval based on semantics of a language rather than plain text.

[0084] FIG. 5 illustrates an exemplary dataset (500) generated using multilingual translation and transliteration in different Indian languages, in accordance with some embodiments of the present disclosure. In FIG. 5, the multilingual dataset (500) comprising a multilingual list of queries associated with an English query is shown. The list of queries include both translated and transliterated query. The multilingual dataset (500) is similar to the multilingual dataset (408) shown in FIG. 4. The queries are stored under different classification (CLS) and are marked by separators (SEP) at the end of each query. For example, without limitations, the sample query shown in FIG. 5 may include the following multilingual variations stored in the database.

Title: “Nirvana Men's Wings Brown T-shirt”

[HINDI]-->. [CLS] नरिवाण मेन ##स वगि ##स ब्राउन टी - शर्ट [SEP]

[BENGALI]--> [CLS] নি ##রভ ##না পুরুষদের উইং ##স ব্রাউন টি- শার্ট [SEP]

[ENGLISH]--> [CLS] Ni ##rvana Men ' s Wings Brown T – shirt [SEP]

[0085] FIGs. 6A and 6B illustrate exemplary query tables (600-A, 600-B) respectively, comprising a variety of queries in various languages, in accordance with some embodiments of the present disclosure. In FIG. 6A, the query table is shown with title row (602) and multilingual queries (604, 606, 608). The title row (602) includes the title “Query: Text” and “Output: comprising fashion items related to the query”. Row (604) illustrates a query in Bengali and how the system has retrieved the correct data based on the query. Similarly, rows (606) and (608) illustrate a query in different languages and the display of the related image.

[0086] In FIG. 6B rows 610, 612, and 614 show queries related to different language and the corresponding output.

[0087] A person of ordinary skill in the art will appreciate that these are mere examples, and in no way, limit the scope of the present disclosure.

[0088] FIG. 7 illustrates an exemplary computer system (700) in which or with which embodiments of the present disclosure may be utilized. As shown in FIG. 7, the computer system (700) may include an external storage device (710), a bus (720), a main memory (730), a read-only memory (740), a mass storage device (750), communication port(s) (760), and a processor (770). A person skilled in the art will appreciate that the computer system (700) may include more than one processor and communication ports. The processor (770) may include various modules associated with embodiments of the present disclosure. The communication port(s) (760) may be any of an RS-232 port for use with a modem-based dialup connection, a 10/100 Ethernet port, a Gigabit or 10 Gigabit port using copper or fiber, a serial port, a parallel port, or other existing or future ports. The communication port(s) (760) may be chosen depending on a network, such a Local Area Network (LAN), Wide Area Network (WAN), or any network to which the computer system (700) connects. The main memory (730) may be random access memory (RAM), or any other dynamic storage device commonly known in the art. The read-only memory (740) may be any static storage device(s) including, but not limited to, a Programmable Read Only Memory (PROM) chips for storing static information e.g., start-up or basic input/output system (BIOS) instructions for the processor (770). The mass storage device (750) may be any current or future mass storage solution, which may be used to store information and/or instructions.

[0089] The bus (720) communicatively couples the processor (770) with the other memory, storage, and communication blocks. The bus (720) can be, e.g. a Peripheral Component Interconnect (PCI) / PCI Extended (PCI-X) bus, Small Computer System Interface (SCSI), universal serial bus (USB), or the like, for connecting expansion cards, drives, and other subsystems as well as other buses, such a front side bus (FSB), which connects the processor (770) to the computer system (700).

[0090] Optionally, operator and administrative interfaces, e.g. a display, keyboard, and a cursor control device, may also be coupled to the bus (720) to support direct operator interaction with the computer system (700). Other operator and administrative interfaces may be provided through network connections connected through the communication port(s) (760). In no way should the aforementioned exemplary computer system (700) limit the scope of the present disclosure.

[0091] While considerable emphasis has been placed herein on the preferred embodiments, it will be appreciated that many embodiments can be made and that many changes can be made in the preferred embodiments without departing from the principles of the disclosure. These and other changes in the preferred embodiments of the disclosure will be apparent to those skilled in the art from the disclosure herein, whereby it is to be distinctly understood that the foregoing descriptive matter to be implemented merely as illustrative of the disclosure and not as limitation.

ADVANTAGES OF THE PRESENT DISCLOSURE

[0092] The present disclosure provides a better retrieval accuracy and diversity compared to Contrastive Language-Image Pre-training (CLIP) and Multilingual Representations for Indian-Languages (MURIL) architecture.

[0093] The present disclosure provides search retrieval for multilingual queries and also handles issues related to translated and transliterated queries.

[0094] The present disclosure applies attention mechanism to improve accuracy of the retrieved results.

[0095] The present disclosure proposes a loss function to improve the diversity of the retrieval results.

[0096] The present disclosure provides label denoising and category extraction. The present disclosure may be used in various applications such as image search engines, content-based image retrieval, multilingual search in Indic setting, zero-shot title/description de-noising, etc.

We Claim:

1. A system (110) for enabling multimodal and multilingual search in a digital platform, said system (110) comprising:

one or more processors (202); and

5 a memory (204) operatively coupled to the one or more processors (202), wherein the memory (204) comprises processor-executable instructions, which on execution, cause the one or more processors (202) to:

encode text data associated with a product in the digital platform in one or more languages;

10 encode an image associated with the product in the digital platform;

and

associate the encoded text data with the encoded image to create a multilinguistic training dataset (408) for generating a text-image encoded matrix for the product in the digital platform.

2. The system (110) as claimed in claim 1, wherein the one or more languages comprise one or more regional languages.

3. The system (110) as claimed in claim 2, wherein the encoded text data comprises at least one of: a translated text and a transliterated text associated with the one or more regional languages.

4. The system (110) as claimed in claim 3, wherein the multilinguistic training dataset (408) comprises at least one of the translated text and the transliterated text corresponding to each of the one or more regional languages associated with the product in the digital platform.

5. The system (110) as claimed in claim 1, wherein the memory (204) comprises processor-executable instructions, which on execution, cause the one or more processors (202) to:

25 receive a text query associated with the product in the digital platform from a computing device (104) associated with a user (102), wherein the text query comprises at least one of: a foreign language query and a regional language query; and

30 transmit one or more details corresponding to the product based on the text-image encoded matrix to the computing device (104).

6. The system (110) as claimed in claim 5, wherein the regional language query comprises a transliterated query represented in foreign language.

7. A method for enabling multimodal and multilingual search in a digital platform, said method comprising:

5 encoding, by one or more processors (202), text data associated with a product in the digital platform in one or more languages;

 encoding, by the one or more processors (202), an image associated with the product in the digital platform;

 associating, by the one or more processors (202), the encoded text data with
10 the encoded image to create a multilinguistic training dataset (408); and

 generating a text-image encoded matrix for the product in the digital platform based on the multilinguistic training dataset (408).

8. The method as claimed in claim 7, wherein the one or more languages comprise one or more regional languages.

15 **9.** The method as claimed in claim 8, wherein the encoded text data comprises at least one of: a translated text and a transliterated text associated with the one or more regional languages.

10. The method as claimed in claim 9, wherein the multilinguistic training dataset (408) comprises at least one of the translated text and the transliterated text corresponding to each
20 of the one or more regional languages associated with the product in the digital platform.

11. The method as claimed in claim 7, comprising:

 receiving, by the one or more processors (202), a text query associated with the product in the digital platform from a computing device (104) associated with a user (102), wherein the text query comprises at least one of: a foreign language query
25 and a regional language query; and

 transmitting, by the one or more processors (202), one or more details corresponding to the product based on the text-image encoded matrix to the computing device (104).

16. A user equipment (UE) (104) for performing multimodal multilingual search in a
30 digital platform, said UE (104) comprising:

 one or more processors communicatively coupled to a system (110), wherein the one or more processors are configured to:

transmit a text query associated with a product in the digital platform to the system (110), wherein the text query comprises at least one of: a foreign language query and a regional language query; and

receive, from the system (110), one or more details corresponding to the product based on the transmitted query and a text-image encoded matrix.

17. A non-transitory computer readable medium comprising a processor with executable instructions, causing the processor to:

encode text data associated with a product in a digital platform in one or more languages;

encode an image associated with the product in the digital platform; and
associate the encoded text data with the encoded image to create a multilinguistic training dataset (408); and

generate a text-image encoded matrix for the product in the digital platform based on the multilinguistic training dataset (408).

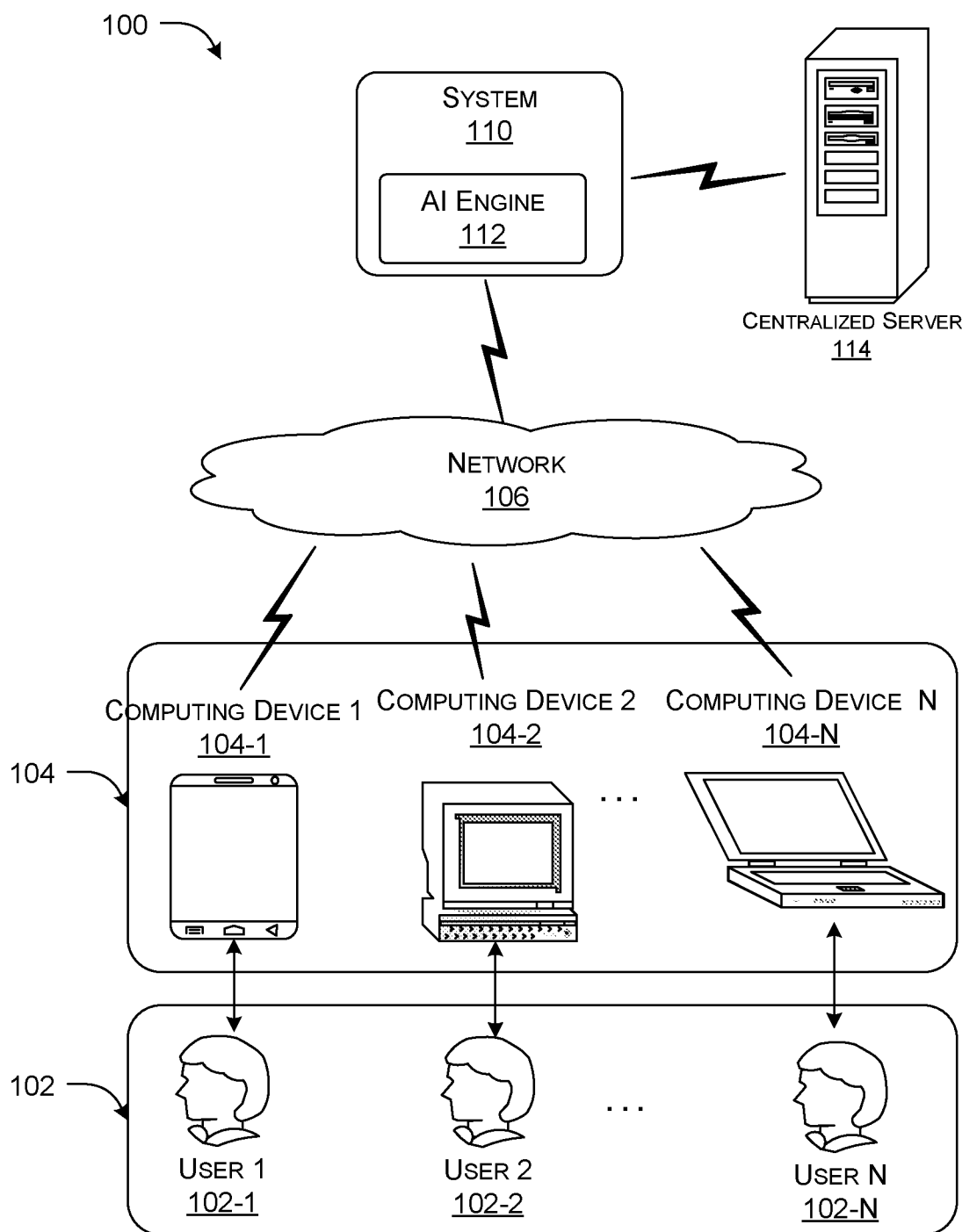


FIG. 1

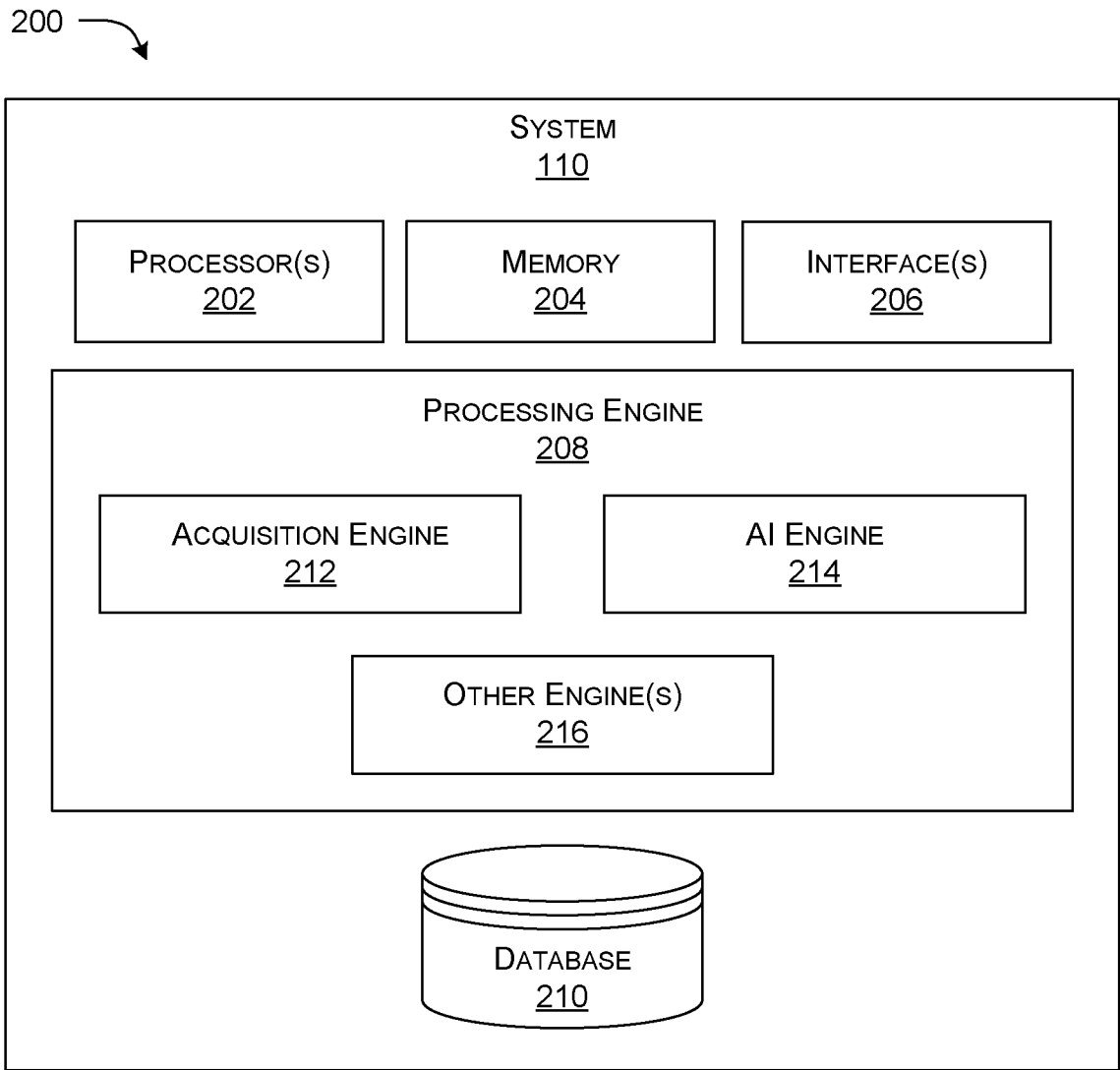


FIG. 2

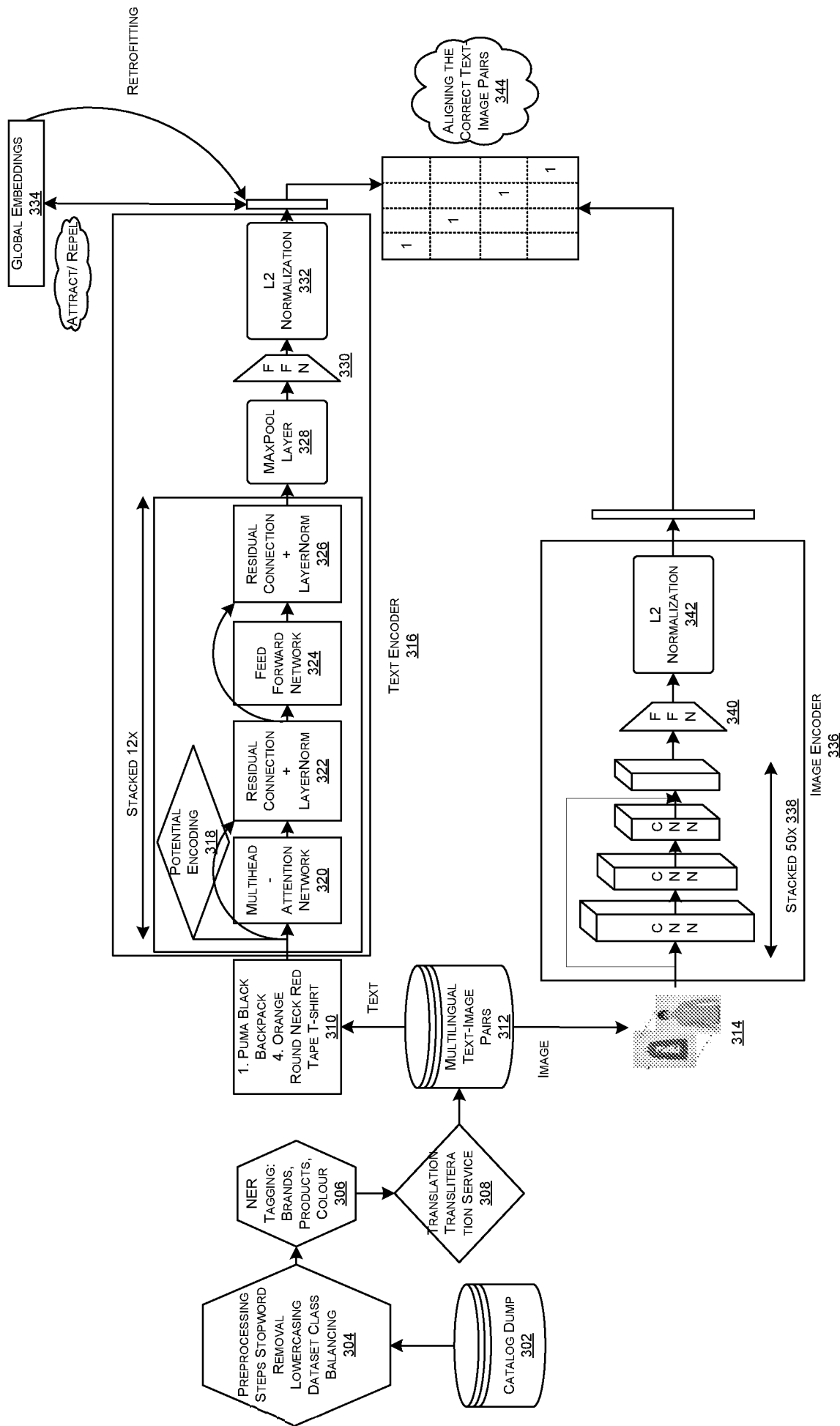


FIG. 3

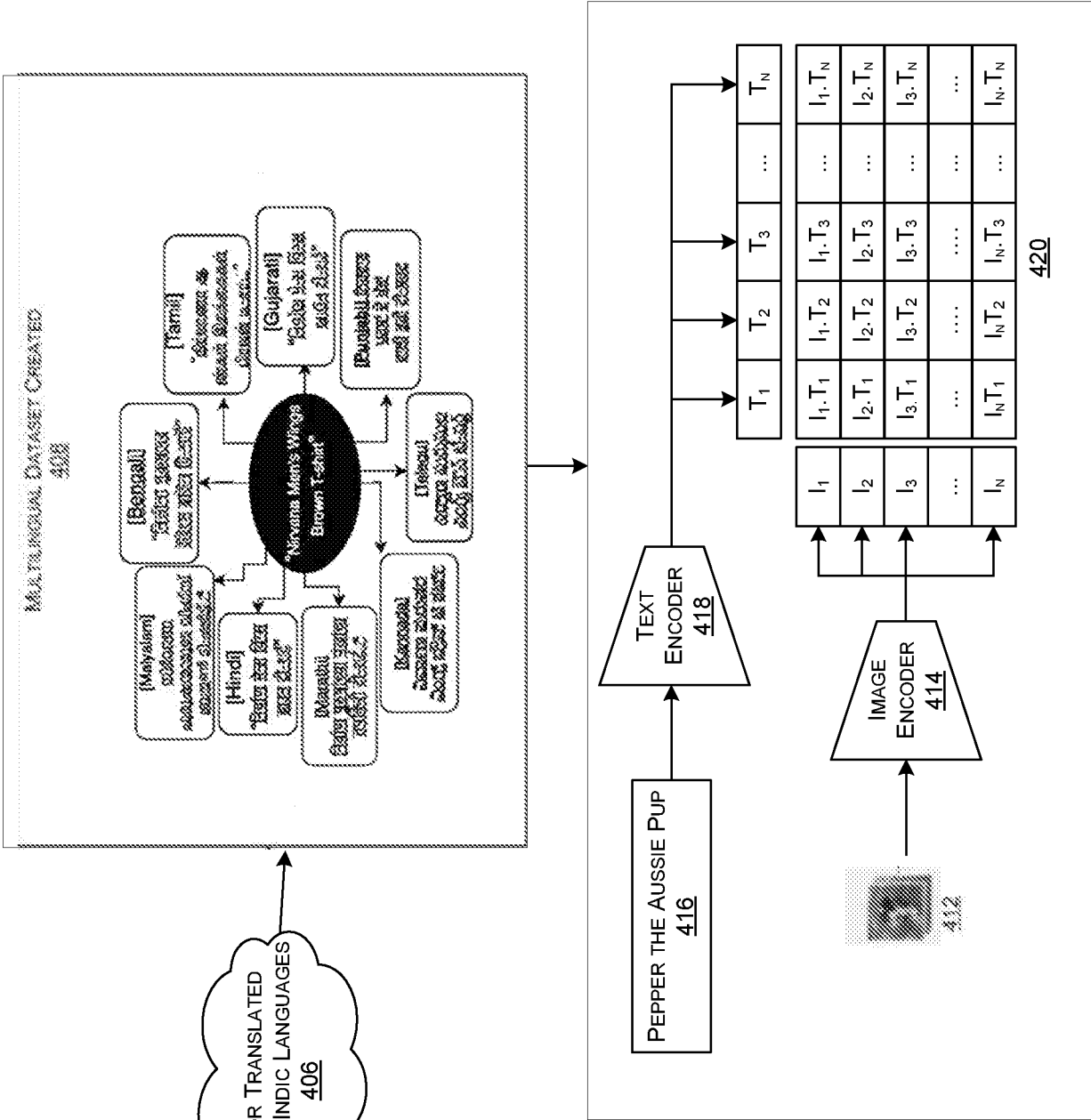


FIG. 4

500 →

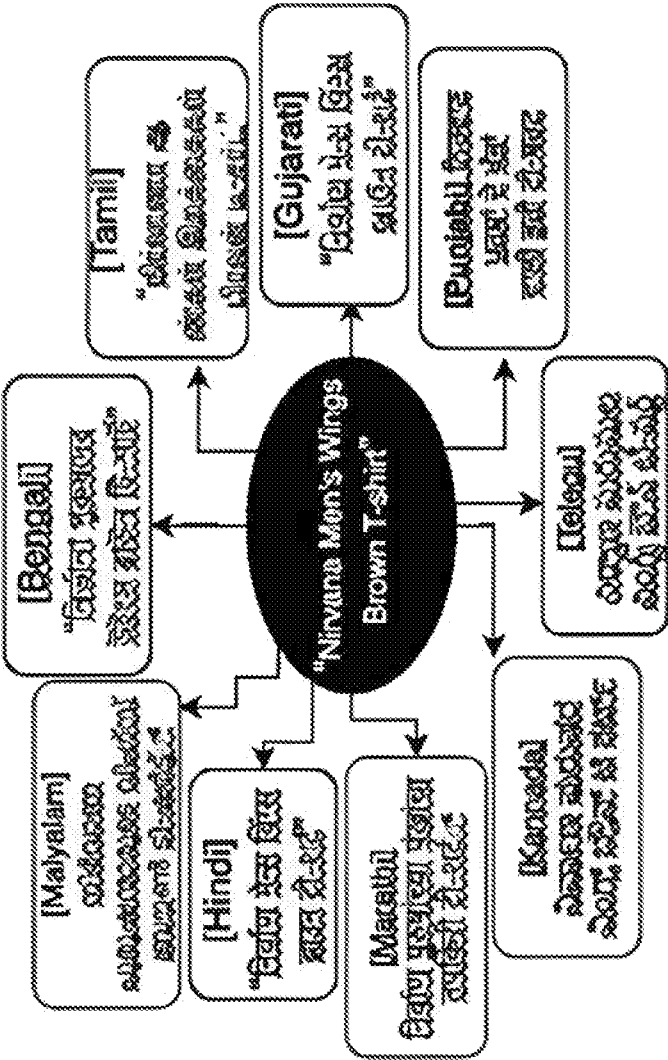


FIG. 5

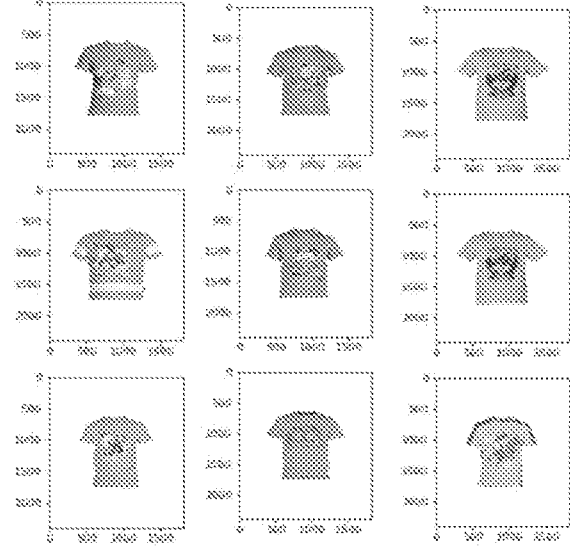
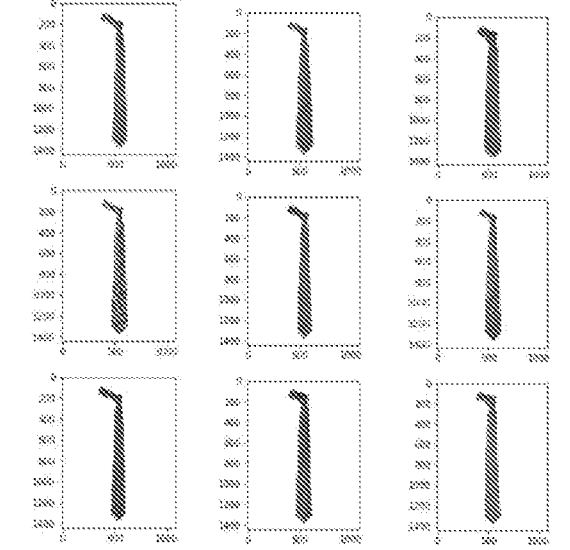
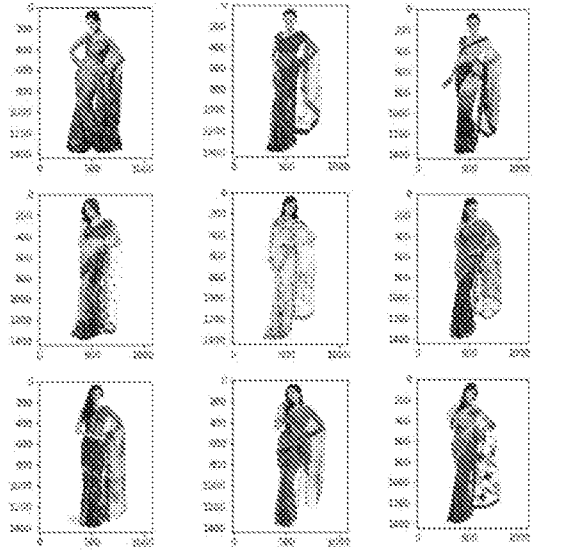
Query : Text <u>602</u>	Output: fashion items related to query
Bengali Query: “ADIDAS কিডস বয়েজ প্রিন্ট করা হলুদ টিশার্ট” <i>[English: “ADIDAS kids boys printed yellow tshirts”]</i> <u>604</u>	
Hindi Query: “रीड आणि टेलर पुरुष लाल टाय” <i>[English: “ani reed tailor red tie for men”]</i> <u>606</u>	
“saree with flowers” <u>608</u>	

FIG. 6A

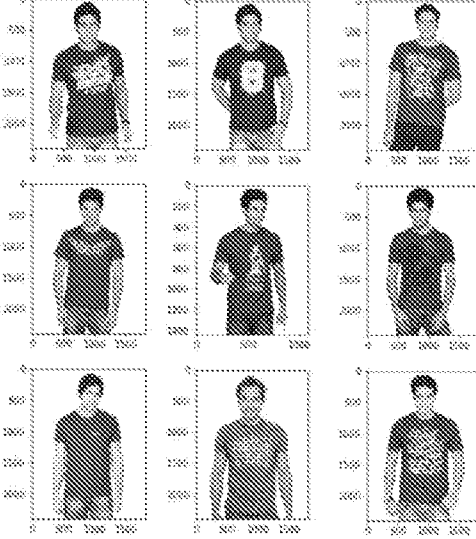
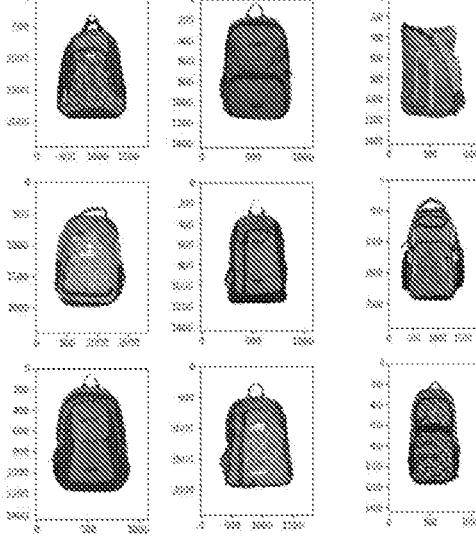
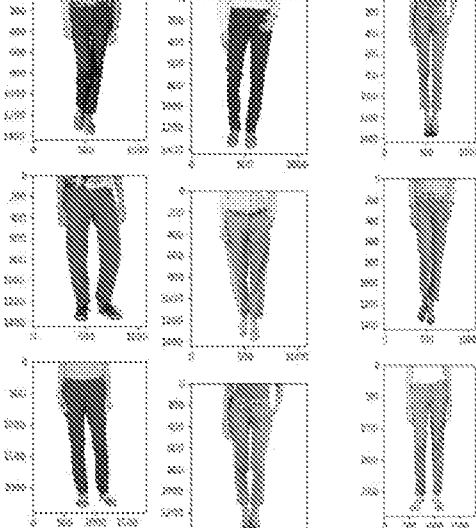
Bengali Query: “পুরুষদের উইংস ব্রাউন টি-শার্ট” [English: “wings brown tshirts for men”] <u>610</u>	
Kannada Query: “ಫುಟ್‌ಲೂಸ್ ಯುನಿಸೆಕ್ಸ್ ಗ್ರೇ ಬೆನ್ನುಹೊರೆಯು” [English: “Footloose Unisex Grey Backpack”] <u>612</u>	
Hindi Query: “फ्रेंच कनेक्शन महिला जैतून पैजामा” [English: “French Connection Women Olive Trousers”] <u>614</u>	

FIG. 6B

SUBSTITUTE SHEET (RULE 26)

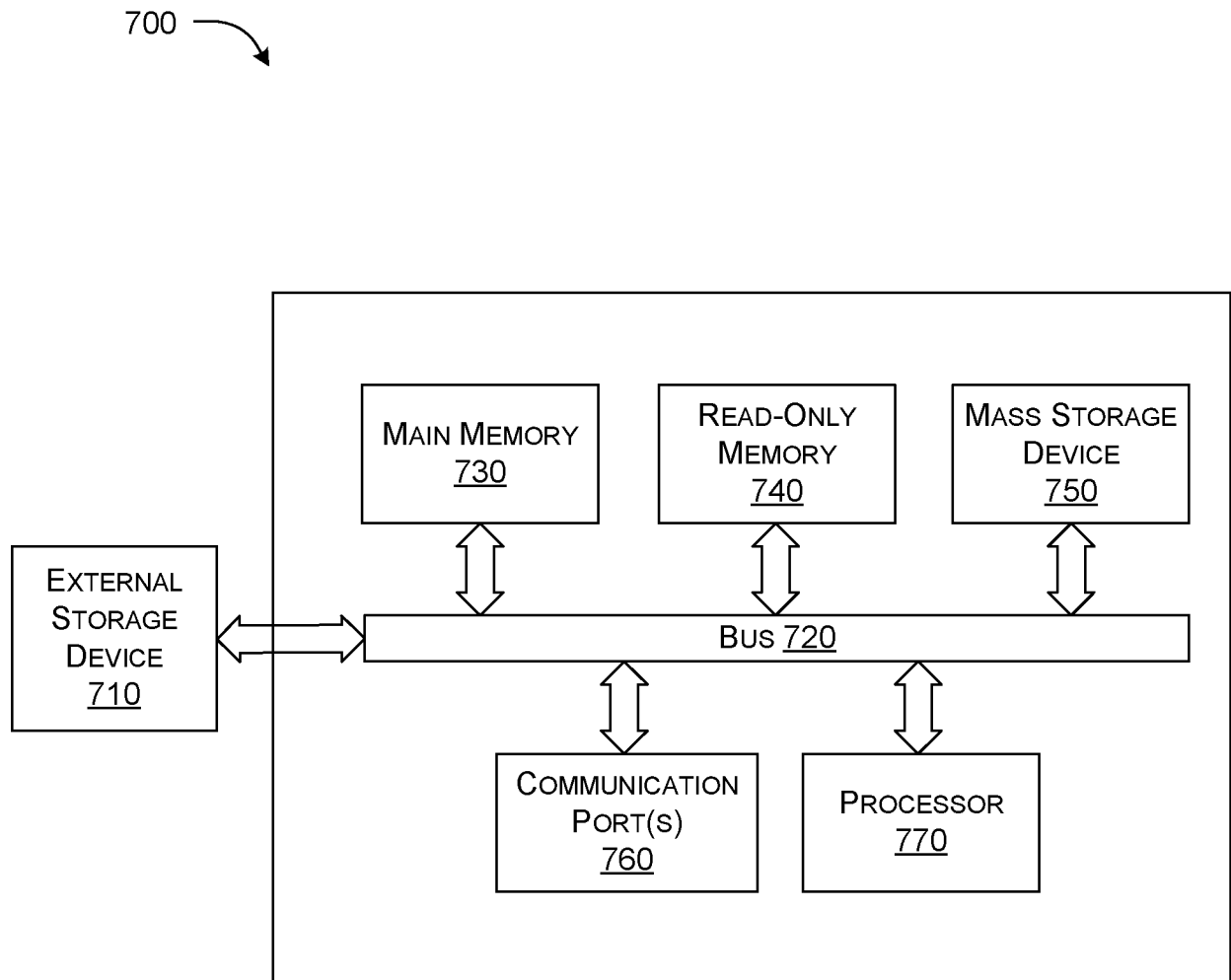


FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB2024/053089

A. CLASSIFICATION OF SUBJECT MATTER

G06F16/24, G06F16/33, G10L15/26 Version=2024.01

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F, G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database consulted during the international search (name of database and, where practicable, search terms used)

PatSeer, IPO Internal Database

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US20180232451A1 (SHUTTERSTOCK INC) 16/08/2018 (16 Aug, 2018) paragraphs [0038]-[0039], [0053], claim 1	1, 7, 16, 17
Y	paragraphs [0038]-[0039], [0053], claim 1 -----	2-6, 8-11
Y	Chakraborty, Neelotpal, et. al. "Language identification from multi-lingual scene text images: a CNN based classifier ensemble approach" Journal of Ambient Intelligence and Humanized Computing 12, 7997-8008 (2021), DOI: "https://doi.org/10.1007/s12652-020-02528-4" Publication date: 19 September 2020 (19-09-2020) [online] refer to sections "Datasets", "CNN architecture"	2-6, 8-11



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

15-07-2024

Date of mailing of the international search report

15-07-2024

Name and mailing address of the ISA/

Indian Patent Office

Plot No.32, Sector 14, Dwarka, New Delhi-110075

Facsimile No.

Authorized officer

Aditya Kashyap

Telephone No. +91-1125300200

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/IB2024/053089

Citation	Pub.Date	Family	Pub.Date
US 20180232451 A1	16-08-2018	WO 2018119200 A1	28-06-2018