Current Technologies and Applications of Digital Image Processing

Wenfeng Zheng
School of Basic Medical Sciences, Peking University, Haidian District, Beijing 100871, P.R.China.
zheng112@hotmail.com

Correspondence should be addressed to Wenfeng Zheng : zheng112@hotmail.com

Article Info
Journal of Biomedical and Sustainable Healthcare Applications (http://anapub.co.ke/journals/jbsha/jbsha.html)
Doi: https://doi.org/10.53759/0088/JBSHA202303002
Received 26 July 2021; Revised from 15 January 2022; Accepted 18 February 2022.

Published by AnaPub Publications

Abstract – There has been significant advances in the field of image and video processing over the past few decades. The term "image processing" is used to describe multiple signal-processing methodologies where images (such as video or picture frames) serves as the input, resulting to another image or a collection of image-related parameters or features. The majority of methodologies to image processing include reducing the picture to a two-dimensional signal and processing it in the same way as any other signal. The term "video processing" on the other hand is used to describe a particular type of signal processing where video files or video streams are utilized as output or input signals. Video recorders, televisions, video codecs, digital versatile, disc players, and other devices all utilize video processing algorithms. This paper provides a survey of the components of Digital Image Processing (DIP) as well as the recent developments in Image Processing technology and DIP applications.

Keywords – Digital Image Processing (DIP), Medical Image Processing (MIP), Non-Photorealistic Rendering (NPR).

I. INTRODUCTION

Digital Image Processing (DIP) [1] is a process where an algorithm is applied to a digital image on a digital computer. DIP is an area of study within digital signal processing that has many merits over its traditional counterpart. It allows additional algorithms to be implemented to the input data, which in turn helps keep noise and distortions from building up as processing continues. For this reason, DIP may be represented by multidimensional systems, since visuals can be defined in more than one dimension. The invention and development of DIP can be traced back to three primary influences: first, the advancement of computer systems; second, the advancement of mathematics (in particular, the establishment and enhancement of discrete mathematics hypothesis); and third, the increasing demand for multiple applications in fields like the environment, industry, medicine, the military and agriculture.

According to Abraham, Morelato, Tahtouh, and Roux [2], most of the methodologies employed in modern DIP were developed in the 1960s for use in fields as diverse as medical imaging, satellite imagery, videophone, wire-photo standard conversion, photograph enhancement, and feature identification at institutions like as Bell Laboratories (BL), the Jet Propulsion Laboratory (JPL), the Massachusetts Institute of Technology (MIT), the University of Maryland (UoM), and a few others. The original goal of image processing was to enhance picture quality. Its intended use was to enhance the way humans seem to the naked eye. It is common practice in the field of image processing to take a low-quality input picture and produce a higher-quality final product. Typical image processing tasks include of improving, restoring, encoding, and compressing images. Taking into consideration the sun's position and the lunar environment, JPL processed the hundreds of images of the moon sent back by the Space Detector Ranger 7 in 1964 by using methodologies such as noise reduction, gradation transformation, geometry correction, etc.

The success of the computer in creating a map of the surface of the moon has had a significant impact. As a result, approximately 100,000 photographs generated by the spacecraft were subjected to advanced image processing, hence producing color maps, topographic maps, and panoramic mosaic of the moon [3]. This then produced significant results and initiated human landing on the moon. However, given the computer advancement of the time, the processing cost was rather significant and expensive. In the 1970s, with the introduction of affordable computers and specialized gear, DIP exploded. As a result, problems like converting between TV standards could be tackled in real time using image processing. General-purpose computer systems began to replace specialized hardware for all but the most computationally intensive tasks as their processing speeds increased [4]. In the 2000s, when powerful computers and signal processors became widely accessible, DIP quickly replaced analog techniques as the standard in the industry.

Digital image processors are developed using the Metal Oxide Semiconductor (MOS) technology, which is dated back in the 1950s' advancement of MOS Field Effect Transistor (MOSFET) at the BL. As a result, new digital semiconductor image sensors like the Charge-Coupled Device (CCD) and the Complementary Metal Oxide Semiconductor (CMOS) sensor emerged. In 1969, Lin et al. [5] developed the CCD and discovered that a small MOS capacitor could be used to hold an electric charge, which they compared to a magnetic bubble. They connected a sufficient voltage to the MOS capacitors in a row so that the charge could be gradually transferred from one to the next. Later, the first digitized video cameras employed the CCD, a semiconductor circuit, to capture and transmit footage for television. Olympus, a Japanese company, developed the NMOS APS (Active Pixel Sensor) in the 1980s. This development was initiated by the condensing of MOS semiconductor device size, with the MOSFET scaling nearing micron and then sub-micron dimensions. In 1980s, Tsutomu Nakamura at Olympus developed NMOS APS before Takayanagi, Nakamura, Fossum, Nagashima, Kunihoro, and Yurimoto [6] discussedthe significant of the CMOS APS.

Nasir Ahmed first presented the Discrete Cosine Transform (DCT) in the 1970s as a lossy compression methodology, which is now standard in digital picture compression [7]. The Joint Photographic Experts Group (JPEG)[8] established a protocol for the compression of images in 1992, basing it on the Discrete Cosine Transform (DCT) [9]. Since JPEG can compress images while keeping their quality intact, it has become the most popular format for storing digital photographs online. Due in significant part to its very efficient DCT compression technology, which led to the widespread adoption of digital images and digital photographs, JPEG pictures numbered in the billions and were being made every day by 2015. There was a paradigm shift in the 1970s, when the MOS technology was widely adopted for use in electronic signal processing. MOS-integrated circuit technology was utilized to develop the first microcontrollers and microprocessors in the 1970s and the first DSP (Digital Signal Processor) chips in the late 1970s. Resultantly, DSP chips have become increasingly popular for use in DIP.

The Discrete-Cosine Transform (DCT), a popular method for reducing images, has inspired numerous businesses to create DSP chips. DCTs are essential for the encoding and decoding of video and audio as well as control signals, multiplexing, signaling, formatting brightness, color differences in color formats like YUV411 and YUV444 and analog-to-digital conversions. DCTs are used in encoding activities such as motion estimations, inter-frame predictions, motion compensations, variable encoding, perceptual weightings, entropy encodings, motion vectors, and quantizations. Decoding processes include inverse operations between RGB, YUV, and YIQ. In most case, DCTs are utilized in HDTV (High-Definition Television) decoder and encoder circuits. British engineer John Housfield of EMI developed the first X-ray CT (Computed Tomoraphy) systems for head diagnostics in the early 1970s. CT nucleus approach reconstructs a cross-sectional picture using a projection of a human head. A CT scanner that could scan the complete body was created by EMI in 1975, enabling precise tomographic imaging of various human body parts. This approach of diagnostics won a Nobel Prize in 1979. The Space Foundation honored medical imaging software used for DIP into the Hall of Fame for Space Technology in the early 1970s.

Although visual and analogue image processing is also viable, DIP is the most sharedtype of image processing, which makes it possible to acquire images that initially produces the input image. A type of video recorder known as digital video operates by employing a digital video signal as opposed to an analog one. This article discusses the components of Digital Image Processing (DIP), recent technologies, and its application. In this article, the phrases camera, camcorder, and video camera are all used interchangeably. The components of DIP will be covered in the sections that follow in this article. Section II focuses on a critical analysis of DIP. Section II provides an analysis of DIP and its components. Section III provides a critical survey of DIP technologies as well as its applications In Section IV, a conclusion regarding the development of DIP is provided.

II. DIGITAL IMAGE PROCESSING

Electron microscope, x-ray machines, still and video cameras, ultrasound and sonar are just some of the physical instruments that create digital images, which are then put to use in fields as diverse as the medicine, arts, commerce, , the military, security, industry civil administration (traffic control), and the sciences. The purpose of any image is to provide a basis for analysis, whether by human or machine. As a result, it is common practice to perform some sort of processing on the raw image before using it. Image enhancement refers to this type of processing, while image analysis refers to processing performed by an observer in order to extract information. Differences in output—images versus scene information—as well as difficulties encountered and strategies used help differentiate between enhancement and analysis. Chemical, optical, and electronic methods have all been used to improve images, while human and computational methodologies have been employed for analysis.

A subfield of electronics, Digital Image Processing (DIP) involves storing an image in a digital memory, processing it using a computer or other digital hardware, and retrieving information from the picture by accessing arrays of tiny numbers, known as pixels, which signify the physical component e.g., scene radiance. As the price of computers continues to drop and their performance increases, DIP is being used more and more often to improve images for human observers or to do autonomous analysis. It is important to remember that images cannot be used as a substitute for direct measurements of physical properties. Instead, it is the outcome of a complex interaction between various physical processes, including the distribution and amplitude of lighting radiations, the scientific knowledge of radiation interplay with matter, which creates the scene of geometry for projecting the transmitted or reflected radiation from 3Ds to 2Ds, and the electrical characteristics of the sensor. There is presently no technique or related theory for extracting scene information of relevance, such as the quality or position of a manufactured product, from a photograph, in contrast to other fields of computer science, such as compiler construction, where algorithms backed by theoretical framework exist for changing a high-level computer language to machine code.

Since most people's own visual system makes information extraction from scenes seem so simple, new users may fail to recognize how difficult it actually is. The visual capabilities of humans are far superior to anything we can create in the near future. In light of this, it is important to remember that human perception is not necessarily indicative of the complexity of a given DIP application. The idea that humans are superior at making judgements while robots excel at quantitative analysis might be considered the first guiding principle. For this reason, DIP works particularly well for tasks like locating and sizing car parts on a conveyor, but struggles with tasks like grading apples and wood (although not impossible). Therefore, digital processing is ideal for image enhancement because it relies heavily on numerical computation but requires little judgment. If extracting usable information from the visual muck wasn't difficult enough, time constraints, which are often strict, make the situation much more difficult. Most commercial uses, for example, must function under severe limitations established by machine cycle durations, but most users likely wouldn't notice if a spreadsheet took extra 300 milliseconds to load rather than 200.

The real-time synthesis of a video stream is also required by numerous applications such as the improvement of ultrasound images, the monitoring of traffic, and the stabilization of camcorder footage. To put the urgency of the situation into perspective, consider that video feeds from normal monochrome video cameras generate roughly 10million pixels each second. As of this writing, the standard desktop computer has less than 50 ms (one hundred nanoseconds) to process every pixel. There is a finite set of actions that can be performed using just 50 instructions. Many applications that use DIP are also subject to stringent budgetary constraints. As a result, we are often confronted with the "engineer's feared triple curse" of having to build something that is both effective and economical. **Fig 1** provides a breakdown of the fundamental steps in DIP.

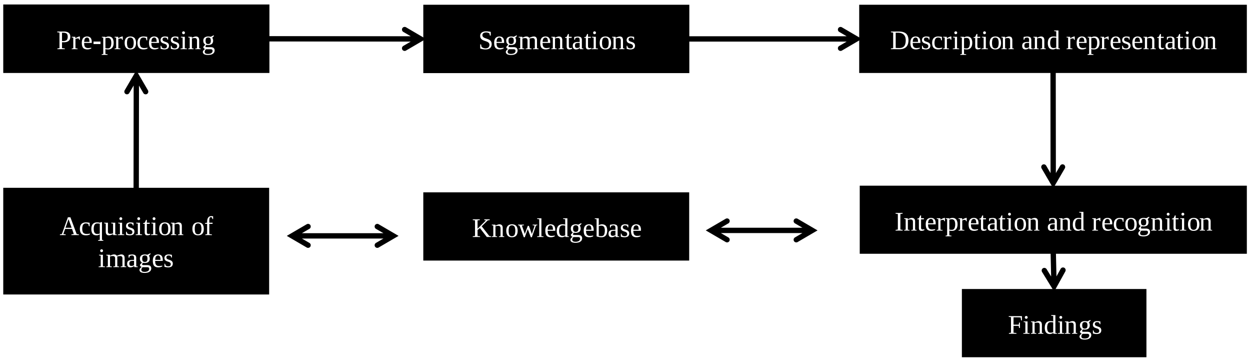

Fig 1. Fundamental steps in DIP

III. APPLICATION OF DIGITAL IMAGE PROCESSING

Many different areas of study and industry make use of DIP; including computer vision, face identification, object recognition, lane departure notification systems, non-photorealistic modeling, medical imaging, and microscopy, remote sensing, morphometric image analysis, and related fields.

Computer Vision

Technology and research have led to the development of computer vision, which allows computers to simulate human eyesight. According to Raval [10], computer vision is the study of the science behind developing artificial systems to extract information from pictures. Video, still images from several cameras, or dimensional information from a clinical scanner are just a few examples of the various possible formats for the picture data. When it comes to building computer vision systems, the ideas and models developed within the field of computer vision are put to the test. Systems for (i) Controlling processes are examples of computer vision applications (e.g., an autonomous vehicle or an industrial robot). (ii) Instance Detection (e.g., for people counting or visual surveillance). (iii) Classifying data (e.g., for indexing image sequences or image databases). (iv) Creating representations of physical entities or whole ecosystems (e.g., industrial inspection, topographical modeling or medical image analysis). (v) Participation (e.g., as the device inputs for human-machine interactions). Additionally, computer vision may be thought of as an adjunct to biological vision rather than a replacement for it. Biological vision is the field that attempts to describe the physiological mechanisms involved in human and animal visual perception. In contrast, the field of computer vision investigates and discusses computer- and hardware-based artificial vision systems. The mutual learning between computer vision and biological vision has been becoming useful. Scene reconstruction, video tracking, event detection, object identification, motion estimation, picture restoration and learning, indexing are all examples of computer vision's specialized sub-fields.

Applications of Computer Vision

Medical computer vision, often known as Medical Image Processing (MIP) [11], is one of the most significant areas of application. In this field, the primary focus is on deducing relevant medical information from patient images. The most common types of imaging data include photographs, X-rays, angiographies, ultrasonic pictures, and tomographic scans. The identification of tumors, arteriosclerosis, or other malignant alterations is all examples of information that may be gleaned from such picture data. Organ size, blood flow, and other similar metrics are also acceptable. When used here, new insights may be gained about the brain's structure or the efficacy of medical therapies, for example, and these can then be used to further medical study. Computer vision also finds use in the industrial sector, where it is referred to as "machine vision" and information is retrieved to aid in the production process.

Automatic inspection for flaws is used in many industries, including quality control, where components or finished products are examined. In addition, data to be extracted by a robotic arm may have their location and orientation measured. One of the most significant uses of computer vision is likely in the military. Identifying enemy troops or vehicles and guiding missiles are two obvious applications. The missile is guided toward an area rather than a specific target by more advanced missile guidance systems, and once there, the missile chooses its target using image data gathered locally. According to "battlefield awareness" and other related concepts in contemporary military theory, a variety of sensors, including image sensors, may give a plethora of information about a clashing situation that can be used to guide strategic choices. In this instance, the processing of the data is automated in order to make it simpler and to aggregate data from many sensors for improved accuracy.

Submersibles, ground vehicles (robots with wheels, trucks or automobiles), Unmanned Aerial Vehicles (UAV)and aircraftare all examples of autonomous vehicles, a relatively recent application field. Vehicles may be entirely autonomous (unmanned) or they can have a driver or pilot supported by computer vision based technologies. In order to identify obstacles and determine their surroundings using simultaneous localization and mapping (SLAM), fully autonomous vehicles often rely on computer vision for navigation. Additionally, it may be put to use in the detection of task-oriented events, e.g., UAV examination of forest fires. Vehicles with obstacle detection and autonomous aircraft landing systems are two types of auxiliary technology. Multiple automakers have shown autonomous driving system prototypes, but the technology is not yet ready for widespread commercialization. There are several types of autonomous military vehicles, from sophisticated missiles to unmanned aerial aircraft used in reconnaissance or missile guiding. Autonomous vehicles equipped with computer vision are already being used in space exploration. One example is NASA's Mars Exploration Rover. Surveillance is another field that may benefit from this technology, as is the aiding of visual effects production for film and television, such as camera tracking (match-moving).

Face Detection

Using a computer algorithm, "face detection" may pinpoint the exact coordinates of a human face in an otherwise random digital picture. It is specifically tuned to recognize human faces, and will disregard other objects or people. Face detection may be thought of as a special example of object-class recognition, where the objective is to identify the sizes and position of all objects in images, which fall into a specified category. According to Rostami, Farajollahi, and Parvin [12], face detection could be considered to be a good example of face localization, where the goal is to pinpoint the exact coordinates of a set of target faces (usually one). When doing face detection, one lacks access to this background data. While earlier face-detection algorithms concentrated on identifying human faces from the front, more recent algorithms have tackled the more challenging and general challenge of identifying faces from several angles. Specifically, it is the ability to recognize faces that have been rotated in either the horizontal (in-plane) or vertical (out-of-plane) planes, or both. The problem of face detection is often implemented in algorithms as a binary pattern classification task.

In other words, a classifier trained on exemplar faces determines whether or not a given portion of an image is a face based on the characteristics extracted from that region. Sliding the window open and shut is a common tactic. In other words, the classifier is used to label all the different sized and shaped (often square or rectangular) parts of a picture as either faces or non-faces (background patterns). When it comes to biometrics, face detection is often used in conjunction with, or as an integral element of, a facial recognition framework. This technology is also employed to aspects such as picture database administration, human-machine interaction, and video surveillance. Autofocus on faces has been a feature of certain newer digital cameras for a while now. Furthermore, in picture slideshows where the Ken Burns effect is used to pan and zoom in on certain areas, face recognition may be a helpful tool for zeroing in on specific people or places of interest.

Feature Detection

The term "feature detection" is used in image processing and machine vision to describe techniques that aim to calculate representations of visual data and carry out local determinations at each image point about the presence or absence of a component of a certain sort at that place. The resultant features will often be discrete points, smooth curves, or interconnected areas that represent subgroups of the image domain. It is not always possible to pin down exactly what a "feature" is, as that will vary depending on the nature of the problem at hand or the context in which it will be used. Because of this, many computer vision algorithms begin with identifying "interesting" parts of an image, which is the definition of a feature. With features serving as the foundation and primary building blocks for subsequent algorithms, the quality of the overall algorithm is often limited by the accuracy of its feature detector.

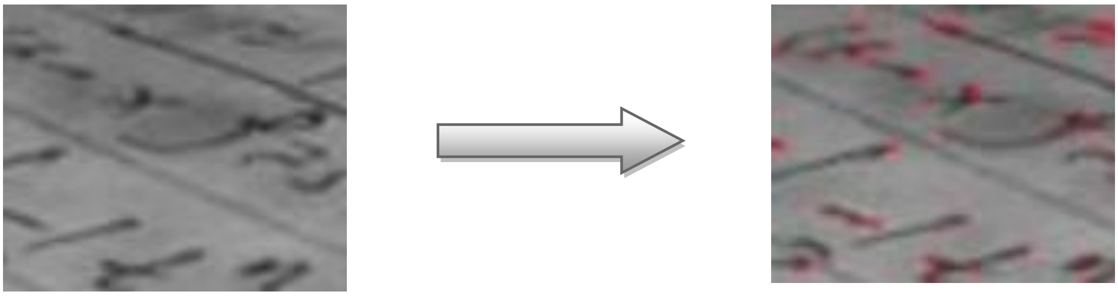

Fig 2. Output of a normal corner recognition methodology

As a result, a feature detector's repeatability—or the degree to which it can reliably detect the same feature in multiple images of the same scene—is an important quality. Detecting features in an image is an elementary process. In other words, it is typically the first step in processing images and analyses every pixel to determine if it contact any features. When incorporated into a larger algorithm, this type of analysis tends to focus solely on the area around the features. For most feature detection methods, preprocessing the input image by applying a Gaussian kernel in a scale-spaced representation and then computing these feature images (described as localized derivative operations) is standard procedure.

It is possible to use a dynamic algorithm to stimulate the feature recognition phase towards only searching relevant regions of an image when performing object recognition is algorithmically expensive and time is of the essence. Due to the prevalence of feature detection as the foundation of many computer vision algorithms, a plethora of different feature detectors have been created. The types of features detected the amount of computational complexity, and the reliability of these methods all vary greatly. In broad strokes, these feature sensors can be classified as one of the following (with some overlap) categories:

Edges

Wherever two different parts of an image meet is where you will find an edge. As a rule, an edge can take on nearly any form and may even feature intersections. In common usage, edges are characterized by clusters of pixels with a large magnitude of gradient. In addition, many widely used algorithms link together a series of high gradient positions to provide a complete definition of edges. Edge characteristics like smoothness, shape, and gradient value are typically bounded by these algorithms. Edges are one-dimensional in their immediate vicinity.

Corners / interest points

Point-like visual elements with a local, two-dimensional structure are called "interest points" or "corners," and the names are often used interchangeably. Named as such because early algorithms used edge detection followed by edge analysis to uncover sharp directional shifts, "Corner" was originally a marketing term (corners). These algorithms were later refined to the point where explicit edge identification was rendered unnecessary, maybe by monitoring the picture gradient for signs of extreme curvature. Then, it was discovered that "corners" were being spotted in areas of the picture (see **Fig. 2**) that were not really corners (For instance, one may be able to pick out a tiny brilliant dot against a black backdrop). Although "interest points" is a more common name for these locations, "corner" has always been the accepted word.

Blobs / interest points or regions of interest

When compared to corners, which are point-like, a blob generates a supplementary explanation of visual models based on regions. Multiple blob sensors could also be considered as an interest point operator since they often produce a favorite point (a point of equilibrium or the peak of an operator's reaction) in their output. If a region of interest in a picture is too smooth for a corner detector to pick up on, a blob detector may be able to help. It is possible to identify corners in a smaller image by first reducing the size of the image. Points that are crisp in the reduced picture but may be blurry in the original will trigger the detector. After this stage, it is hard to distinguish a corner detector from a blob detector. Including a suitable idea of scale may fix this difference to a significant degree. However, the Determinant of Hessians (DoH) and Laplacian of Gaussians (LoG) blob sensors are also discussed in the corner detection articles because to their response qualities to various picture structures at varying scales.

Ridges

The idea of ridges is an intuitive method of describing lengthy objects. As a further refinement of a medial axis, ridge descriptors produced from grey-level picture may be thought of as a dividing line. The local ridge width connected with every ridge point may be thought of as an additional feature of a ridge that could be visualized as a 1D curve signifying a symmetry axis. The algorithmic difficulty of extracting ridge component from wide categories of grey-level pictures is higher than that of extracting corner-, blob, or edge-, features. Despite this, route recovery in aerial photos and blood vessel detection in medical imaging often use the descriptor of ridges.

Lane Departure Warning Model

On motorways and arterial highways, lane departure warning systems are installed to alert drivers when their vehicles begin to veer out of their lanes without their having signaled a turn. In Europe, Iteris's lane departure warning system for Mercedes Actros commercial vehicles was the first of its kind to go into mass production. The technology was introduced in the year 2000, and since then it has been standard on most European-market vehicles. The Iterismodel first appeared on Freightliners Truck in North America in 2002. All these setups use a rumble strip to alert the driver if the car begins to stray from its lane unintentionally. No alerts are issued when a turn signal is activated. In the transportation sector, lane departure warning systems are becoming more successful since they combine preventative and risk reporting.

Viewnyx uses video analysis and other video-based technology, as described by Mudric, Cuk, Janicijevic, Nedeljkovic, and García-Ramos [13], to help fleets reduce their liability expenses associated with driving. First, we must deal with the most common causes of accidents, which are human mistake behind the wheel, distractions, and fatigue. Second, we empower Safety Managers to proactively advise and train employees to reduce risky behaviors by disseminating data and tools for assessing and prioritizing driver and fleet risk. Numerous fleets across North America have used Lookout as their solution of choice. There are two primary categories of lane-departure warning systems: I those that simply alert the driver to the fact that the vehicle is leaving the lane and (ii) those that, if the driver does nothing, take corrective action to keep the car in the path.

Non-Photorealistic Rendering

The field of Non-Photorealistic Rendering (NPR) in computer graphics is concerned with facilitating a broad range of expressive approaches in digital art. According to Wegen, Döllner, Wagner, Limberger, Richter, and Trapp [14], computer graphics have always been centered on photorealism, but NPR takes its cues from a wider range of creative practices, including painting, sketching, technical representation, and animated cartoons. Aside from the typical uses in media like movies and video games, NPR has also been used in the fields of architectural illustrations and experimental animations. Modern uses of this technology include Cel-shaded animation. As an alternative to the more realistic approach taken by standard computer graphics, NPR is a subfield of computer graphics that aims to open the door to a broad range of expressive approaches in digital art. NPR is influenced by other forms of visual art outside animation. Cel-shaded animation (sometimes called "toon" shading) is a kind of NPR that has featured in movies, computer games, computer vision, structural illustration, and experimentation animation. There was a session at SIGGRAPH 1990 called "Non Photo Realistic Rendering," which is often credited as being the origin of the phrase NPR.

There havebeen some critics to the term: First, "photorealism" is defined differently by academics in the field of computer graphics (for more, see "photorealistic rendering") and by visual artistes. These artistes are the major users of the NPR methods, therefore the term refers to a certain style of painting that attempts to recreate the distorted and hyper-reflective qualities of a camera lens. However, in the field of computer graphics, the term "photoreal" refers to a picture that is so realistic that it fools the human eye. In reality, "non-photorealism" is the term used by graphics experts to describe the visual distortions used by photorealist artists.Second, it is not always easy to describe something by what it isn't. Non-elephant biology or non-geometric mathematics are fictitious analogies that may be used instead. Researchers at NPR have suggested that they believe the phrase will be phased out in favor of the more inclusive "computer graphics," with "photorealistic graphics" denoting "classic" computer graphics. Third, many methods for making supposedly "non-photorealistic" graphics are not really rendering methods at all. This refers to modeling or post-processing methods. It is particularly unpleasant for conference organizers because sketch-based modeling approaches cannot be included under the topic of "image-based rendering," despite the fact that the latter is increasingly used to describe the former.

Possible re-brandings were discussed by Jha, Shukla, Ghosh, Khisti, and Dubey [15] during the first symposium on non-photorealistic animations and visualization. The terms "expressive graphics," "artistic rendering," "non-realistic graphics," "art-based rendering," and "psychographics" were all thrown about as possibilities. Different academic articles on the problem have used a variety of words, but "non-photorealistic" has caught on. In 2000, the Association for Computing Machinery (ACM) hosted the first technical symposium focused only on NPR known as the Symposium on NPAR (Non-Photorealistic Animation and Rendering). On odd-numbered years, NPAR is held in conjunction with the AAFF (Annecy Animated Film Festival). For the first time in 2007, NPAR was held on years with an arbitrary amount with ACM SIGGRAPH.

NPR in 3Ds is the standard for video games and movies. The end result of using this method is often a 3D framework, which has been altered from the input framework to potentially reflect a distinct aesthetic form. In many cases, the only difference between the original and the replica is the surface material. NPR effects may now be performed to the rasterized picture before it is presented on the screen thanks to shaders and the widespread availability of programmable GPUs. Most NPR methods for 3D geometry are geared on flattening the image. Both cel shading and Gooch shading are examples of NPR methods for 3D graphics. Several methods, including occluding outlines and suggestive contours, may be used to create stylized outlines and lines from 3D objects. The most efficient methodological graphics for technical communications do not have to be photo-realistic to serve their purpose of improving readability. Exploded view diagrams and other non-photorealistic representations are very helpful for illustrating the arrangement of components in a sophisticated system.

A video or still picture is the usual starting point for a 2D NPR system. While most 2D NPR is used for creative reasons, such as data visualization, the result is often an artistic interpretation of the input imagery. The aesthetic depiction of images and videos has historically been centered on heuristic algorithms, which try to replicate the placements of brush strokes on a digital canvas (also known as image stylization). Paul Haeberli's "Painting by Numbers" presentation at SIGGRAPH 1990 was one of the first examples of 2D NPR [16]. This (as well as related interactive methodologies) provides users with a blank canvas where they can "paint" with a cursor, revealing a stylized representation of the picture as they go. For emulating brush strokes of varying widths in various parts of a picture, this is quite helpful.

As a result, beginning in the late 1990s, simple image processing techniques like statistical moments, or gradient operators were used to digitalize this procedure and reduce the amount of human involvement required (even if the user may still exercise some degree of creative agency by adjusting the algorithms' settings). To apply 2D NPR to video format for the first time, as indicated in the movie “What Dreams May Come's moving paintings” in 1998, required the assistance of automation. In the 2000s, Ryabchikov and St. Petersburg Federal Research Center of the Russian Academy of Sciences [17] utilized the operators of computer vision such as image segmentation and image salience to create more advanced methods of picture abstraction, which in turn drove stroke placement. Image analogy, an algorithm for stylizing images that may learn to emulate the manner of an existing piece of art, was heavily influenced by machine learning at this time. Image stylization has been given new life thanks to deep learning, specifically Neural Style Transfer (NST) algorithms that could mimic broad varieties of creative models with just a single sample. These algorithms provide the basis of smartphone applications that may do similar tasks, for instance. Prisma Similar to the aforementioned stylization approaches, a group of 2D NPR techniques focus on the imitation of creative mediums. Ink diffusion across various papers and watercolor pigment dispersion in water are both simulated using these techniques.

Medical Image Processing

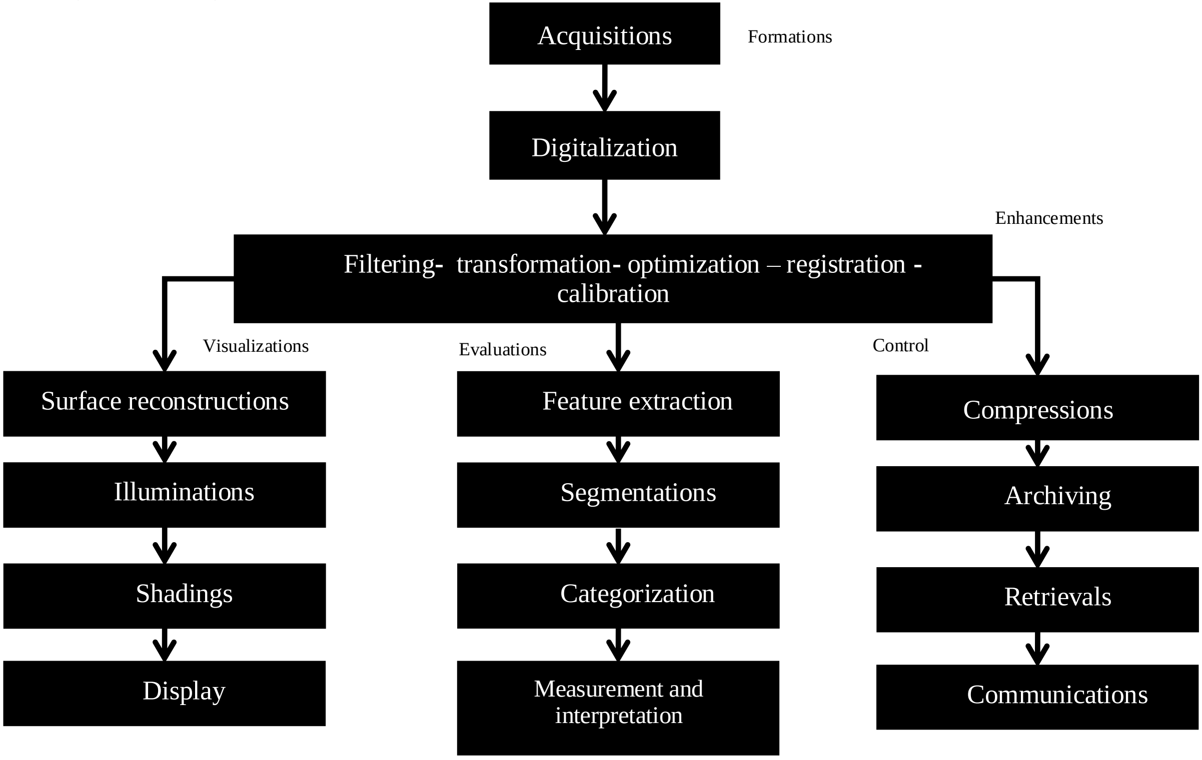

Fig 3. Image processing system

When employed in clinical settings (medical procedures meant to disclose, diagnose, or investigate illness), medical imaging refers to the method and process of producing images of the human body (body parts and functions) for those same objectives (such as the analysis of typical physiology and anatomy). Radiology (in the wider perspective), nuclear medicine, medical photography, endoscopy, microscopy, medical thermography, and investigative radiological science are are all included in this field of study (e.g. for human pathological investigation). Even though electroencephalography (EEG), electrocardiography (EKG), and magnetoencephalography (MEG), were not originally developed with the intention of producing images, the data they collect is highly suitable for representation as maps (i.e., integrating positional data), so these approaches could be visualized as a subset of medical imaging.

Medical imaging has come a long way since the first of x-rays in 1895, according to a research by Bry, Saenz, Pappas, Kalaitzakis, Papanikolaou, and Rasmussen [18]. Due to the proliferation of direct digital imaging equipment, the healthcare industry has seen a rise in the importance of DIP. Analogue imaging modalities, e.g., radiography and endoscopy, have recently has modern sensors added to them in order to better compete with originally digital technologies such as magnetic resonance imaging (MRI) and CT. Individual "picture elements," or "pixels" (an abbreviation derived from the terms "picture" and "element"), are what make up digital pictures, and they each have their own unique value for lightness and color saturation. With the use of the Picture Archiving and Communication Systems (PACS) and the Digital Imaging and Communications in Medicine (DICOM) protocol, medical images may be processed, reviewed, and shared quickly and reliably across multiple of devices and locations. The medical field may now benefit from the full breadth of DIP due to the widespread use of the digital imaging processing system illustrated in **Fig. 3.** Commonly, when someone refers to "medical image processing," they imply the service of processing digital images for use in medicine. There are five main categories in MIP (see **Table 1**):

Table 1. Representation of MIP categories

MIP categories	
Image formation	Image formation consists of photographing the subject, editing the photo, and creating a digital image matrix.
Image visualization	Image visualization is used to describe any process that modifies this matrix to provide a more useful representation of a picture.
Image analysis	Image analysis entails a number of processes, all of which contribute to the final product, which may be utilized for either concrete or abstract purposes, such as quantitative measures. These processes need for highly abstracted a-priori information about the pictures' nature and content to be included into the algorithms. As a result, image analysis is a highly specialized process, and it is unusual for newly discovered algorithms to find immediate use in other contexts.
Image management	Image management involves data storage, data transfer, data archiving, and data access (retrieval). Compression methods are used since a basic grayscale radiograph in its uncompressed state might take up a lot of space (perhaps several gigabytes). Similarly, telemedicine techniques may be thought of as a kind of image management.
Image enhancement	Image enhancement, refers to either human or automated procedures that may be employed without a-priori data about the accurate contents of images, in contrast to high-level processing, also known as image analysis. No matter what is shown in a picture, the results of this sort of algorithm are the same.

A technique is considered low-level or high-level for processing an image regardless of how complicated the algorithm is, how challenging it is to implement, or how much CPU time is needed. The difference lies, rather, in the level of abstractions of the a-priori information used to make the judgment. The information about pictures is simplified from a concrete (iconic) to an abstract (symbolic) description. Low-level image processing techniques are those that work directly with the underlying data at the pixel, edge, or texture levels. There are several different approaches to high-level image processing, e.g., scene, object, region, and texture levels. Better modelling of a-priori information may help humans attain the necessary level of abstraction.

These criteria highlight a specific challenge in the advanced processing of medical images: due to its complex nature, medical a-priori information is difficult to articulate in a way, which it could be integrated readily and directly into digital algorithms of image processing. The semantic gap refers to the disjuncture between the physician's high-level cognitive assessment of a diagnostic picture and the low-level representation of an image in a computer program, where individual pixels serve as the building blocks (low level). It is challenging to bridge the semantic gap in the medical field for four key reasons identified in **Table 2**.

Table 2. Key reasons to bridge the semantic gap in the medical field

Low image quality	Most diagnostic and therapeutic imaging methods are hazardous to human health. This means that photos are captured with minimal energy or dosage, resulting in a poor signal to noise ratio.
Heterogeneity of images	Organs and other bodily components are visible in medical photographs. These objects' form, size, and structural components may change significantly not just from patients to patients (inter-subject variation), but also across various views of the same patient and identical views of the same patient at different times, even if obtained with the same modalities and using a consistent acquisition process (intra-subject variation). That is to say, there is room for variation in biological structures between and among individuals. This means that a priori knowledge cannot be expressed universally.

Unknown delineation of objects	The therapeutically or diagnostically significant item is often represented throughout the whole picture, making it impossible to segment biological entities from their surroundings. Even if recognizable things may be seen in a medical picture, it might be difficult to segment them since the image only shows a partial or distorted version of the object's form. Consequently, the highest degree of abstraction often possible for medically-related things is the texturing level.
Robustness of algorithms	Besides these intrinsic features, which make high-level processing of medical pictures difficult, the medical field has unique criteria for the robustness and dependability of clinical processes and, when employed routinely, the algorithms of image processing. Automatic medical picture analysis shouldn't consistently result in inaccurate diagnoses or treatment recommendations. This necessitates the immediate dismissal of any picture that cannot be processed properly. As a result, it is essential that accurate assessments be made of all non-rejected images.

There has been a shift in the way medical images are processed during the last several decades (see **Fig 4**). Accelerating processing speed was a priority in the 1980s as studies focused on the creation of digital pictures, the proper administration of image data, and image improvement and visualization. Since the 1990s, assessment of medical image interpretation algorithms and their applications in areas such as registration, fragmentation, classification, and accurate measurements have been prioritized. The future will favor the promotion of MIP's incorporation into the daily routines of practicing doctors. Further uniformity is needed when high-level image processing is integrated into the diagnostic and therapeutic processes. This means that MIP will likely continue to be a dynamic area of study and practical application in the healthcare, scientific research and clinical education sectors.

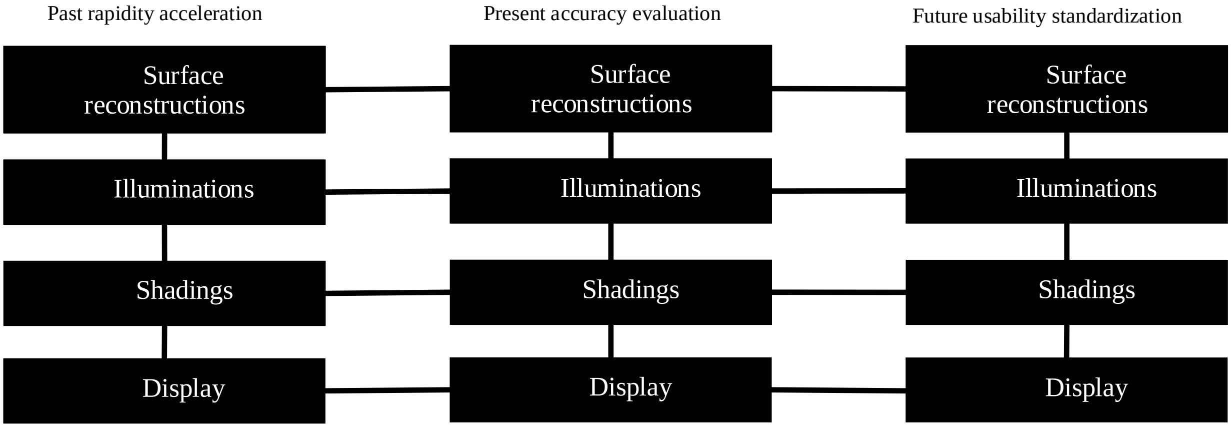

Fig 4. Paradigm of MIP

Magnetic Resonance Imaging

In order to create images of the body, a MRI device or scanner, initially called a "Nuclear Magnetic Resonance (NMR) imaging scanner," employs strong magnets to polarize and excite hydrogen nuclei (single protons) in water molecules identified in human tissues. An extremely powerful stationary static field (magnetic field) is used to radicalize the hydrogen and helium; a weak time-varying (on a 1 kHz) field (the gradient field) is used for spatial encoding; and a weak signal field is used to manipulate the hydrogen nucleus to develop detectable signals, collected via the RF antenna.

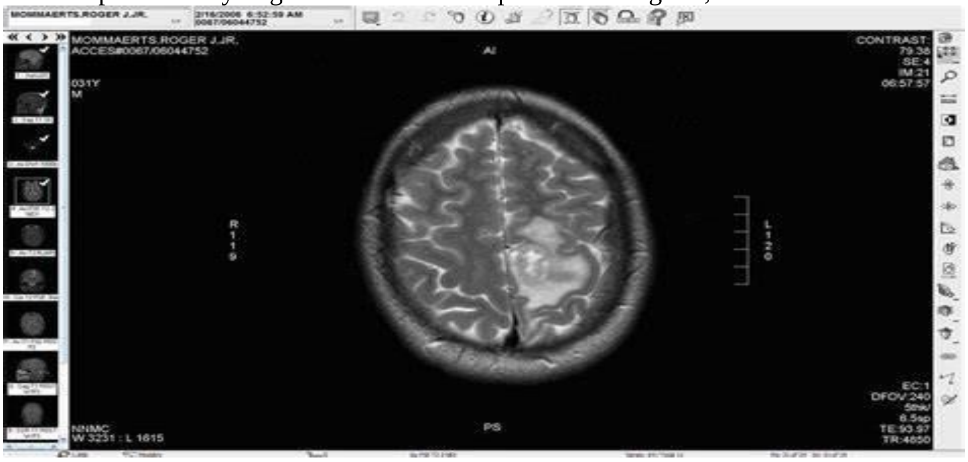

Fig 5. Brain MRI representations

The Magnetic Resonance Imaging (MRI) technique is categorized as a tomographic imaging method because, like computed tomography, it creates a 2D picture of a thin "slice" of the body as seen in **Fig 5**. (CT). The tomographic, single-slice concept may be extended by modern MRI equipment to produce pictures as a 3D block. Magnetic resonance imaging (MRI), which does not use onizing radiation, has no such health hazards. There is no limit on the amount of scans a person may get using an MRI, in contrast to CT and X-ray, since there are no known long-term effects of exposure to dynamic static regions. Tissue heating as a result of exposures to RF area of the available of implanted devices, e.g., pacemakers, is, nevertheless, connected with known health hazards. The instrument's architecture and the scanning processes put a lid on these potential dangers. Images acquired with CT and MRI seem quite different from one another because the two methods are sensitive to very distinct tissue characteristics. Since X-rays must be stopped by thick tissue in CT in order to obtain a picture, the quality of the image will suffer while studying soft tissues. Although any nucleus with a general nuclear spin could be employed in MRI, the protons of hydrogen atoms is still the most popular, particularly in the medical field as a result to its high signal return and signal abundance. The availability of this nucleus in the water molecule is what makes the MRI soft-tissue contrast so impressive.

Microscope Image Processing

Using DIP methods to modify, analyze, and display microscopical pictures is what we call "microscope image processing," and it's a very large field. This kind of processing is increasingly prevalent in many different areas, including medical, biology, medical research, health screenings, material science, etc. More than one microscope maker now routinely includes hardware and software components intended to facilitate communication with an external image processing system.

Morphological Image Processing

Mathematical Morphology (MM) [19] refers to a discipline that utilizes concepts from set theory, lattice hypothesis, topologies, and random variables to analyze and manipulate geometrical shapes. Surface meshes, solids, graphs, and many other spatial systems are all suitable for using MM; however it is most typically used with digital pictures. MM may be used to describe topological and geometrical continuous-space perspectives including connectedness, convexity, size, form and geodesic distance on both continuous and discrete spaces. In addition to its use in morphological analysis, MM serves as the basis for the collection of operators known as mechanical image processing, which modifies pictures in accordance with the aforementioned characteristics. Aside from its use with binary images, MM has also been used to grayscale algorithms and pictures. Subsequently, MM was extended to full lattices, which is now generally recognized as its theoretical basis.

Remote Sensing

The term "remote sensing" [20] refers to the practice of gathering data on an item or phenomena from a distance, without coming into direct touch with it. This may be done on a local or global scale (e.g., by satellite, aircraft, buoy, ship, or spacecraft). To put it simply, remote sensing is the technique of acquiring data about an object or area from a distance using technological instruments. Examples of remote sensing include earth assertions or weather satellite gathering platforms, atmosphere and ocean monitoring weather buoys, surveillance of parolee with ultrasound detection systems, X-ray, PET, space probes and MRI. Currently, the phrase is used to distinguish the area of imaging from others, such as medical imaging, and to describe the use of image sensor technologies, such as those used in airplanes and spacecraft and in neurobiology.

Remote sensing may be divided into two categories: passive and active. Passive remote sensors rely on the detection of naturally occurring radiation given off or reflected by the viewed item or environment. Passive sensors typically detect radiation from reflected sunlight. Video photography, charge-coupled devices, Infrared, and radiometers are all examples of passive remote sensors. On the other hand, active sensors involve outward emissions of energy to scan regions and objects, followed by the detection and measurement of radiation rebounded or transmitted and reflected from the target. One type of active remote sensing is radio detection and ranging (RADAR), which uses the time lag between an object's emission and subsequent return to determine its position, altitude, velocity, and heading. By using remote sensing, information may be gathered about regions that are otherwise too hazardous or inaccessible to send a team there. Some of the many uses of remote sensing technology include tracking the rate of deforestation in places like the Amazon Basin, gauging the impact of global warming on polar ice caps, and measuring ocean and coastline depths. During the cold war, the military used remote sensing to gather information on potentially volatile border regions.

IV. CONCLUSION

The study of image and video processing has recently increased, making it a popular field. In this paper, "image processing" has been referred to a broad range of signal-processing methodologies where images (such as video or picture frames) serving as outputs or inputs might be a set of parameters or single image attributes associated with images. The bulk of methods for processing images include splitting the image into two dimensions and treating it like any other signal. Digital Image Processing (DIP) is a branch of electronics that entails saving an image in a digital memory, processing it with a computer or other digital equipment, and extracting information from the image by accessing a collection of minuscule numbers, known as pixels, that represent a physical characteristic like scene radiance. DIP is used in many different fields of research and business, such as computer vision, face recognition, object recognition, lane departure warning systems, non-photorealistic modeling, medical imaging, and microscopy, morphometric image analysis, and remote sensing. By eliminating the need for time-consuming and resource-intensive data collecting on the ground, remote sensing helps to preserve natural environments and prevents the destruction of valuable assets. Together with data collected through ground-based detection or larger scale aerial analysis, the data collected and sent from orbital platforms allows scientists to keep tabs on natural long- and short-term phenomena like El Nio. Earth science data may also be put to use in a variety of other contexts, including national security and the collecting of information from the air, the ground, and a distance along international borders, as well as in agricultural domains like land utilization and conservation.

Data Availability

No data was used to support this study.

Conflicts of Interests

The author(s) declare(s) that they have no conflicts of interest.

Funding

No funding was received to assist with the preparation of this manuscript.

Ethics Approval and Consent to Participate

Not applicable.

Competing Interests

There are no competing interests.

References

[1]. M. A. Adoul, M. Abed, and A. Belouchrani, “Time-frequency readability enhancement of compact support kernel-based distributions using image post-processing: Application to instantaneous frequency estimation of M-ary frequency shift keying signals,” Digit. Signal Process., vol. 127, no. 103535, p. 103535, 2022.

[2]. J. Abraham, R. Ng, M. Morelato, M. Tahtouh, and C. Roux, “Automatically classifying crime scene images using machine learning methodologies,” Forensic Science International: Digital Investigation, vol. 39, no. 301273, p. 301273, 2021.

[3]. L.-A. Tran, S. Moon, and D.-C. Park, “A novel encoder-decoder network with guided transmission map for single image dehazing,” ProcediaComput. Sci., vol. 204, pp. 682–689, 2022.

[4]. Computer.org. [Online]. Available: https://www.computer.org/about/cs-history/1970s. [Accessed: 01-Jan-2023].

[5]. W. Lin et al., “Analysis of ionizing and department damage mechanism in proton-irradiation-induced scientific charge-coupled device,” WuliXuebao, vol. 64, no. 2, p. 024220, 2015.

[6]. I. Takayanagi, J. Nakamura, E. R. Fossum, K. Nagashima, T. Kunihoro, and H. Yurimoto, “Dark current reduction in stacked-type CMOS-APS for charged particle imaging,” IEEE Trans. Electron Devices, vol. 50, no. 1, pp. 70–76, 2003.

[7]. E. Gul and A. N. Toprak, “Contourlet and discrete cosine transform based quality guaranteed robust image watermarking method using artificial bee colony algorithm,” Expert Syst. Appl., vol. 212, no. 118730, p. 118730, 2023.

[8]. H. Perolla, “Comparative study on MATLAB based joint photographic experts group image size reduction using shearlet and wavelet packet transform for X-ray images with potential hospital data storage applications,” Rev. Gest. Inov. Tecnol., vol. 11, no. 2, pp. 1312–1323, 2021.

[9]. Y. Zhou, L. Hong, and X. Zhu, “Document image watermarking scheme with adaptive positioning based on discrete wavelet transform and discrete cosine transform,” in 2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP), 2022.

[10]. V. Raval, “Design and development of a computer vision algorithm and tool for currency recognition in Indian vernacular languages for visually challenged people,” ELCVIA Electron. Lett. Comput. Vis. Image Anal., vol. 18, no. 2, p. 4, 2020.

[11]. R. Yang, J. Yu, J. Yin, K. Liu, and S. Xu, “A dense R-CNN multi-target instance segmentation model and its application in medical image processing,” IET Image Process., vol. 16, no. 9, pp. 2495–2505, 2022.

[12]. M. Rostami, A. Farajollahi, and H. Parvin, “Deep learning-based face detection and recognition on drones,” J. Ambient Intell. Humaniz. Comput., 2022.

[13]. M. Mudric, I. Cuk, D. Janicijevic, A. Nedeljkovic, and A. García-Ramos, “Feasibility of a modern video-based technology for assessing the reaction time during specific karate kumite situations,” Int. J. Perform. Anal. Sport, vol. 20, no. 4, pp. 620–630, 2020.

[14]. O. Wegen, J. Döllner, R. Wagner, D. Limberger, R. Richter, and M. Trapp, “Non-photorealistic rendering of 3D point clouds for cartographic visualization,” Abstr. Int. Cartogr. Assoc., vol. 5, pp. 1–2, 2022.

[15]. A. K. Jha, P. Shukla, P. Ghosh, P. Khisti, and A. Dubey, “Visualization and measurement of natural convection boundary layer by particle image velocimetry,” J. Flow Vis. Image Process., vol. 30, no. 2, pp. 1–22, 2023.

[16]. P. Haeberli, “Paint by numbers: abstract image representations,” in Proceedings of the 17th annual conference on Computer graphics and interactive techniques - SIGGRAPH ’90, 1990.

[17]. I. A. Ryabchikov and St. Petersburg Federal Research Center of the Russian Academy of Sciences, “A method for automatic recognition of deviant behavior of people based on the integration of computer vision and knowledge management technologies to support decision-making by operators of video monitoring systems,” Analysis and data processing systems, no. 3, pp. 21–36, 2022.

[18]. V. Bry, D. Saenz, E. Pappas, G. Kalaitzakis, N. Papanikolaou, and K. Rasmussen, “End to end comparison of surface-guided imaging versus stereoscopic X-rays for the SRS treatment of multiple metastases with a single isocenter using 3D anthropomorphic gel phantoms,” J. Appl. Clin. Med. Phys., vol. 23, no. 5, p. e13576, 2022.

[19]. O. Lézoray, “Mathematical morphology based on stochastic permutation orderings,” Math. Morphol. - Theory Appl., vol. 5, no. 1, pp. 43–69, 2021.

[20]. L. Tian, J. Tian, J. Wang, X. Wang, and W. Li, “A novel remote sensing index for brine shrimp (Artemia) slick detection in salt lakes,” Remote Sens. Environ., vol. 286, no. 113428, p. 113428, 2023.