Computer Vision: A Timely Opportunity

Takudzwa Fadziso

Institute of Lifelong Learning and Development Studies, Chinhoyi University of Technology, **ZIMBABWE**

*Corresponding Contact:
Email: takudzwafadziso@gmail.com

ABSTRACT

Different approaches to the study of computer vision have been taken into consideration. It begins with the collection of raw data and advances to methodologies and ideas that combine digital images, pattern recognition, machine learning, and computer graphics to produce new products, as well as new products themselves. In order to get crucial information, students can analyze images and videos to interpret events or descriptions, as well as discern patterns in the landscape, using computer vision. It makes use of a multi-spectrum application domain strategy in conjunction with a large amount of data in order to achieve its goals. In recent years, technological developments in computer vision have paved the way for the creation of novel agricultural applications. Precision yield forecasts for fruit and vegetable crops are particularly critical for improving harvesting, marketing, and logistics planning and execution. When a bridge is under stress or has a high volume of traffic, the geographical and temporal information provided by cars on the bridge reflects this. It is proposed to design a methodology for information gathering and dissemination by utilizing computer vision technology, which recognizes various items tracking and picture calibration via a quick regional neural convolution network, and a quick regional neural convolution network (Faster R-CNN). When dealing with small fish populations, it can be difficult to objectively assess the differences in behavior between individuals. The behavior of fish in aquaculture tanks has been studied with the use of a computer vision system that has been built in order to quantify these types of observations. Contained traffic load data is essential for bridge statistical analysis, security evaluation, and maintenance planning. This is particularly true for heavy trucks. From retail to agriculture, and across all industries, computer vision is having a big impact on organizations of all sizes and in all sectors. When a human eye is required to assess the situation, the significance of this becomes even more apparent. This paper provides information about computer vision technology, including short algorithms, issues, opportunities, and applications for computer vision in a range of fields in the year 2021, as well as information on computer vision in general. Information about computer vision applications in many fields is also included for the year 2021.

Key words:
Computerized navigation, concurrent computing, computer vision movement

9/2/2021 **Source of Support:** None, **No Conflict of Interest:** Declared

Introduction

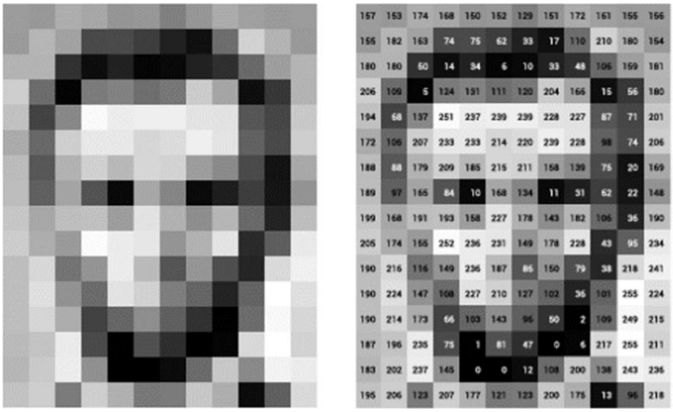
Computer vision is an area of computer science that is concerned with the development of digital systems that process, evaluate, and make sense in the same manner that people use visual information in their daily lives. It is also known as computer vision technology (images or videos). The concept of computer vision is built on computers that train you to evaluate and understand an image at the pixel level, rather than at a higher level of abstraction, as opposed to higher levels of abstraction. Simply put, machines attempt to retrieve, handle, and interpret visual information through the use of software algorithms that are distinct to each machine, which is a very simple technical explanation. In the following list, you will find some instances of jobs where computer vision technology are widely employed. The system scans the visual input and detects the object as a photo or video according to the category that has been defined in the settings. Objects are divided into categories. For example, the algorithm can recognize a dog in an image despite the fact that there are many other elements present. The visual content of an item identifier is analyzed by the system, which assesses whether or not a certain thing is portrayed in a photograph or a video clip. For example, the algorithm can recognize a specific dog in a photograph among a big number of dogs in a large number of photographs. Video systems are being processed. It is the process of recognizing and tracking the movement of an item (or objects) in response to search parameters. Object tracking can be defined as follows: In addition to face recognition, deep picture analysis, which allows for visual searches such as those conducted by Google Images, and biometric identification methods, computer vision can be used to address more sophisticated problems such as object recognition, object tracking, and object tracking (Bynagari & Amin, 2019).

Literature Review

Computer extracts the information from any object by using its algorithm and optical sensors (Matiacevich et al., 2013). Computer vision is a branch of AI and has expanded into human visualization. To put it another way, computer vision allows the intelligent machine to see, process and act according to algorithms (Khan et al., 2021). Feature extraction, image acquisition, pre-processing, high-level processing, detection/segmentation and decision-making are the common frameworks being used in computer vision (Patel et al., 2012). This framework has two categories e.g. pixel optimization and 3D morphological analysis. 3D morphological leads to pattern recognition and image processing, whereas pixel optimization includes the pixel morphology in which internal elements of vector function and structural analysis are studied (Nandakumar et al., 2016). This tactic ought to be practiced on comparatively large datasets which covers the many layers of geometrical composition. Computing algorithms that are used to extract the quantitative information comprehends the overall complex clusters of colour (Bynagari, 2020). Artificial intelligence tactics and morphological analysis can better be performed by using computing algorithms. They can be combined to perform even the most difficult jobs (Bakirtzis et al., 2016).

Working Methodology of Computer Vision

It has been observed that when it comes to computer vision technology, it has a tendency to mimic the operation of the human brain. The question is, how our brain deals with the problem of visual object recognition in the first place. Some people believe that our brains use models to decode specific items in our environment, which is one of the most widely held beliefs. This method is used in the development of computer-based vision systems, which are becoming increasingly popular (Ahmed et al., 2021). It's difficult to assess the effectiveness of algorithms used in creation in bringing our own internal mental processes closer together because we don't fully understand how images integrate the intellectual and the visual. At its most fundamental level, computer vision is concerned with the recognition of patterns. In order to train an artificial intelligence system to recognize visual information, the most effective method is to feed it images—a large number of images, if possible—as well as thousands or millions of labels, and then subject them to a variety of software technologies or algorithms that enable the computer to recognize patterns in every element associated with the labels, such as colors, shapes, and patterns in motion.



157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Figure 1: Pixel data diagram

In the case of a million photographs of cats, for example, the computer's algorithms can analyze the colors of the photographs, the shapes and distances between shapes, the bordering of items, and other characteristics in order to identify and quantify a profile of what the term "cat" represents. Following completion, the computer will (in theory) be able to use its knowledge to find new unlabeled photographs of cats if it is provided with new unlabeled photographs of cats to search through. An easy gray-scale image buffer is shown in the following code, and it is used to save our Abraham Lincoln photograph. The luminosity of each pixel is represented by a single 8-bit value with a value range ranging from 0 (black) to 255 (white), which is represented in the above Figure 1.

```
{157, 153, 174, 168, 150, 152, 129, 151, 172, 161, 155, 156,
 155, 182, 163,  74,  75,  62,  33,  17, 110, 210, 180, 154,
 180, 180,  50,  14,  34,   6,  10,  33,  48, 106, 159, 181,
 206, 109,   5, 124, 131, 111, 120, 204, 166,  15,  56, 180,
 194,  68, 137, 251, 237, 239, 239, 228, 227,  87,  71, 201,
 172, 105, 207, 233, 233, 214, 220, 239, 228,  98,  74, 206,
 188,  88, 179, 209, 185, 215, 211, 158, 139,  75,  20, 169,
 189,  97, 165,  84,  10, 168, 134,  11,  31,  62,  22, 148,
 199, 168, 191, 193, 158, 227, 178, 143, 182, 106,  36, 190,
 205, 174, 155, 252, 236, 231, 149, 178, 228,  43,  95, 234,
 190, 216, 116, 149, 236, 187,  86, 150,  79,  38, 218, 241,
 190, 224, 147, 108, 227, 210, 127, 102,  36, 101, 255, 224,
 190, 214, 173,  66, 103, 143,  96,  50,   2, 109, 249, 215,
 187, 196, 235,  75,   1,  81,  47,   0,   6, 217, 255, 211,
 183, 202, 237, 145,   0,   0,  12, 108, 200, 138, 243, 236,
 195, 206, 123, 207, 177, 121, 123, 200, 175,  13,  96, 218};
```

Figure 2: Unsigned chars

In reality, pixel values are almost always stored in a one-dimensional array at the hardware level, unless otherwise specified. Using the image above as an example, the data has been stored in a manner that is similar to the following long list of unsigned hexadecimal characters (Figure 2).

Despite the fact that when displayed, image data appears to be two-dimensional, this method of storing image data can be used, contrary to your expectations. But this is because computer memory is simply a linear list of address spaces that is constantly growing in size, as opposed to other types of storage.

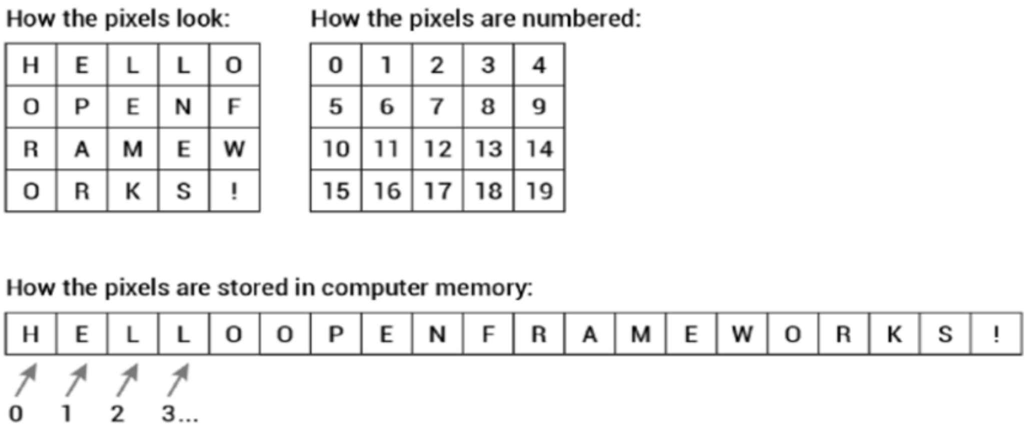


Figure 3: How pixels are stored

Bring the picture back up and think about whether it would be beneficial to include a colored version of it this time. At this point, things are becoming more complicated to understand. The majority of computers read color on the same scale from 0–255 for three different values: red, green, and blue (in the vast majority of cases) (RGB). Each pixel now contains three additional values, which the computer can use to store the information in it. If President Lincoln were to colorize his portrait, the result would be 12 x 16 x 3, which equals 576 different combinations of letters and numbers.

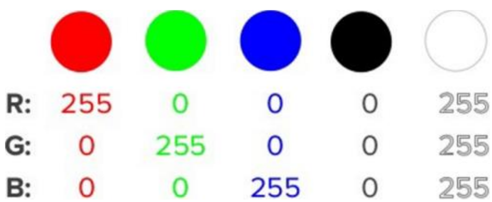


Figure 4: Creation of colors with RGB

When storing an image, a large amount of memory is required, and when iterating an algorithm, a large number of pixels are required. However, in order to train a model with meaningful accuracy, particularly when talking about Deep Learning, you would typically need tens of thousands of images, and the more you can do it, the better the model will perform (Bynagari, 2018).

Image Classification

Convolutional neural networks (CNNs) are the most widely used neural network architecture for image classification today, accounting for more than half of all neural network architectures (CNNs). For CNNs, a typical application would be to feed network images into the network and to classify the data coming from the network, both of which are common. To begin training a CNN with "scanner" input, which does not attempt to parse all of the training data at the same time, is common practice in machine learning. When a layer with 10,000 nodes attempts to enter a 100 x 100-pixel image into another layer with 10,000 nodes, you do not want this to happen!

As an alternative, you can create an input scanning layer of 10 x 10 pixels that receives the first ten ten pixels of the image and feeds them into the computer. Immediately after receiving this input, the next 10 x 10 pixels are added to the screen by shifting the scanner to the right one pixel on the screen. The term "sliding windows" is used to describe this particular technique.

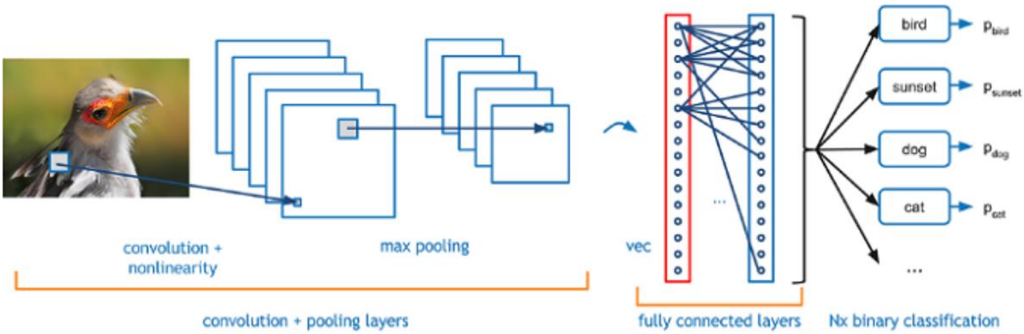


Figure 5: Convolutional neural layers

When feeding data into convolutional layers, convolutional layers are used rather than normal layers to ensure that the data is properly distributed. Each node only has to deal with cells that are close to it in terms of distance. Convolutional layers have a tendency to become smaller as they become deeper, which is due to the easily divisible input factors, which is a primary reason for this tendency. In addition to convolutional layers, pooling layers are also commonly used to enhance the performance of the convolutional layer system. The technique of swimming is a method of filtering details: one common method of swimming is max pooling, which involves taking a set of pixels (for example, two x two pixels) and passing the attributes with the greatest number of attributes onto each pixel in turn.

In recent years, the Image Net dataset, which contains approximately 1.2 million high-resolution images, has been used to train the vast majority of pictorial classification techniques, including supervised learning. The test images are not displayed with an initial annotation because no segmentation or labels have been applied to them; instead, the algorithms must generate labels that specify the objects visible in the images in the absence of segmentation or labels.

Object Detection

When the sliding window technique is used for image classification and localization, as is the case here, a CNN must be applied to a large number of different image crops. Because CNN categorizes every crop as either an object or a background, we must apply CNN to a large number of locations and scales, which is both time-consuming and computationally costly.

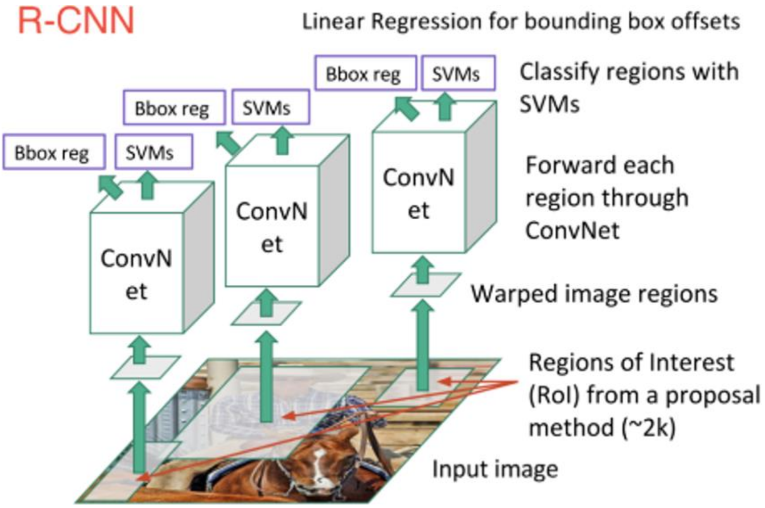


Figure 6: Linear regression for bounding box offsets

For this reason, researchers who studied neural networks proposed that regions where "blobby" image regions that are likely to contain objects are found should be avoided and that these regions should be utilized instead.

Runnable distance is only a few hundred meters. The R-CNN model was the first to be put into practice, in 1995. (Region-based Convolutional Neural Network). Selective Search is used to search for potential objects in an input image in the R-CNN. This results in approximately 2,000 region proposals per image, which is used to train the R-CNN. After that, in addition to each of these regional proposals, we also run a CNN.com broadcast.

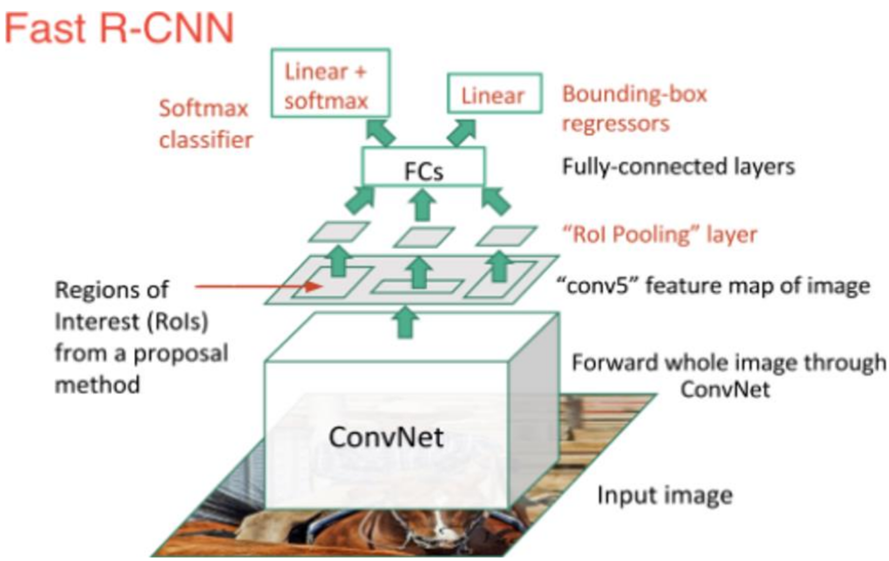


Figure 7: Fast R-CNN

The problem of object detection was transformed into a problem of image classification, in essence, by our approach. While it does have its advantages, it also has some disadvantages, including a slow training process, a large amount of disk space requirement, and a slow deduction process. In terms of structure, Fast R-CNN is a direct descendent of the R-CNN algorithm. As a result, detection speed is increased in two ways: (1) it performs functional extraction prior to regions proposing only one CNN over the entire image, and (2) it replenishes SVM with a softmax layer to extend the neural network for predictions rather than creating a new model.

However, even though the selective search algorithm generates proposals for regions in a short period of time, it takes a long time to finish a search.

In order to achieve this result (RPN). When determining "where" to look, the RPN is utilized in order to reduce the amount of time spent on a computer during the entire deduction process. When the RPN scans each location, it does so quickly and efficiently, allowing it to determine whether or not additional processing in a particular region is required. In order to do so, it generates k bounding box offers with two scores, each of which represents the probability that the object will appear or will not appear at each of the k locations in question.

Once we have received proposals from our region, we feed them directly into a Fast R-CNN for further consideration and analysis. Before we get to the classification layer for softmax and the bounding reverberator, which is the final layer in our chain, we have to go through several fully connected layers.

Compared to slower R-CNNs, faster R-CNNs were capable of achieving significantly higher speeds and precision. Note that, despite the fact that future models have proven to be extremely useful in increasing detection speeds, only a few models have been able to significantly outperform Faster R-CNN in terms of overall performance.

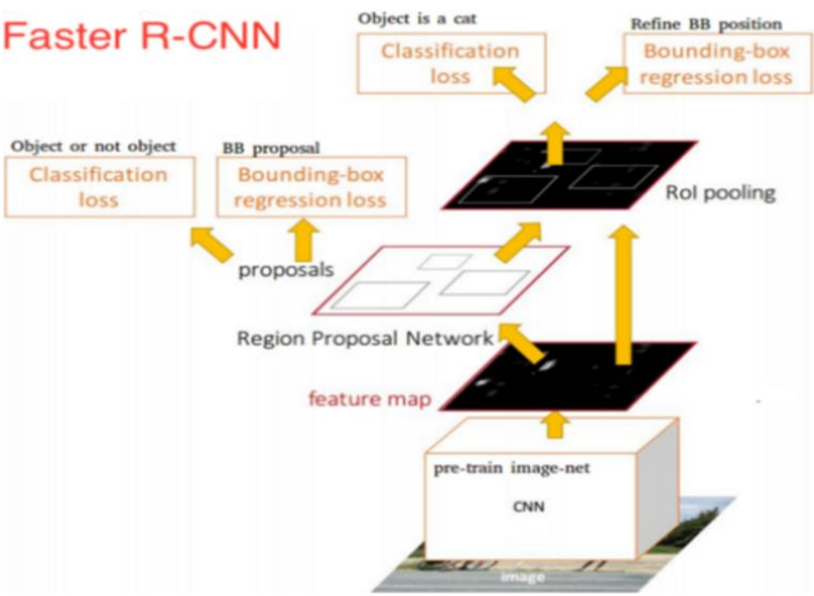


Figure 8: Faster R-CNN

Over the past few years, there has been a significant shift in the object detection industry away from slower, less efficient detection systems toward faster, more efficient detection systems. There were several approaches that used shared computations on a single image, such as You. All of which were examples of approaches that used shared computations on a single image, such as You Look Only Once (YOLO). This distinguishes them from the expensive sub-networks used by the three R-CNN techniques, which are used in these approaches. The primary reason for these trends is that they do not focus on their respective sub-problems separately.

Object Tracking

When a specific object or a group of objects of interest is being tracked across the landscape of a given scene, this process is referred to as object tracking, or tracking. In terms of observer models, methods for object tracking can be divided into two categories: generative observer models and discriminative observer models. The apparent features of an object are described by a generative model in the case of the generative method, which reduces the reconstruction error in object searching techniques such as principal component analysis (PCA).

As a result of its dependable performance, it is becoming increasingly popular as the primary tracking technique for aircraft. Deep learning is included in this category because it is a discriminatory method that is also known as tracking by detection and is therefore included in this category. This is accomplished by identifying candidate objects for all frameworks and tracking them down.. Object recognition is accomplished through deep learning, which is applied to the candidates that have been identified. The SAE (stacked autoencoders) and neural network convolution (NNC) models are the two most fundamental types of network models that can be used in machine learning applications (CNN).

Deep Learning Tracker is the most popular deep network used for tracking tasks with SAE, and it is the most widely used because it provides both offline pre-training and online networking capabilities. The following is a description of how the procedure should be carried out:

- In this study, large-scale natural image datasets were used for uncontrolled pre-training of the stacked self-encoder, which was then used to represent general objects in the real world. Auto-encoder stacked denoising can achieve a more robust feature expression capacity than other methods by introducing input photo noise and reconstructing the original images after they have been processed.
- Combining the pre-trained network coding component with the classification network enables you to use positive and negative samples from the original setup to refine the network, which can then distinguish between the current object and the background.
- While being used as a motion model, DLT employs particulate filters to create potential patch candidates for the currently displayed frame. To select the patch that will be the subject of the classification, the classification network calculates the probability scores for each of these patches, which represent how confident it is in its classification. The patch with the highest probability score is then selected as being the object of the classification.
- When updating the model, DLT uses the limited threshold path, which is more efficient.

CNN has become widely used. It is also referred to as a deep model. Generalized convolutional neural networks (CNNs) can be trained to perform both classification and tracking tasks. CNN tracking algorithms such as the fully revolutionary network tracker (FCNT) and the multi-domain CNN tracking algorithms are two examples of CNN tracking algorithms (MD Net).

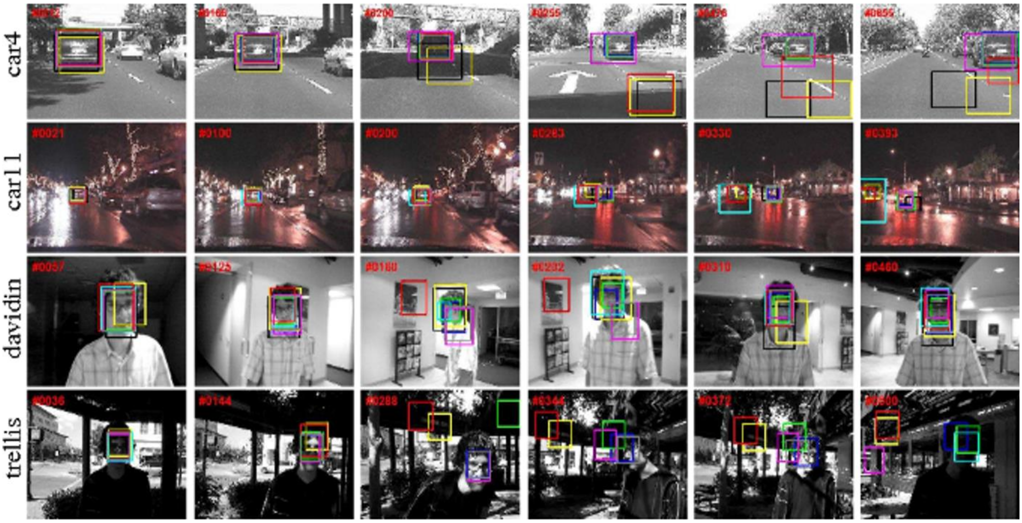


Figure 9: Object tracking by various categories

With the help of the VGG model, which is a pre-trained ImageNet model, FCNT can successfully analyze and employ feature maps, and it makes the following observations about the model:

- CNN maps can be used to track down and locate a specific individual.
- In many cases, CNN maps are either too noisy or unrelated to be able to distinguish a specific object from its immediate surroundings.
- Higher layers encode semantics on categories of objects, whereas lower layers encode intra-class variation by encoding more discriminatory functions, thereby capturing intra-class variation.
- Higher layers encode semantics on categories of objects, whereas lower layers encode intra-class variation by encoding more discriminatory functions, thereby capturing intra-class variation.

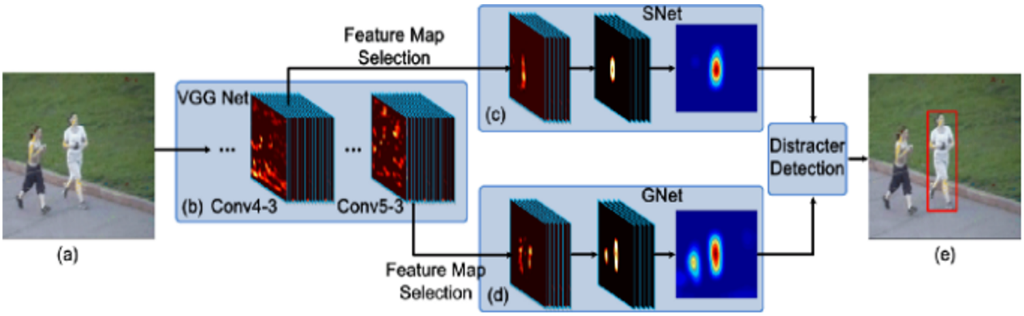


Figure 10: Pipeline of FCNT

When creating both networks, it is necessary to initialize them with the bounding box for them to function properly. Also included is the cropping and propagation of a region of interest (ROI) centered on an object located in the previous frame, which will be used in the following frames. Once this is accomplished, two predictive heat maps are received by the classification device through the use of SNet and GNet, and the tracker determines which of the two heat maps will be used to generate the final tracking result based on whether or not there are any distractors in the final tracking result. An illustration of the FCNT pipeline is shown in the diagram below.

MD Net, in contrast to FCNT, tracks movement throughout all video sequences, rather than just a few of them. This is a departure from traditional tracking methods and practices, as the networks mentioned above make use of irrelevant image information to reduce the training requirements for tracking data. A multidomain concept is proposed by MD Net to distinguish between the object of a class in this video and the background of another video if the object and background of one video are the same in both videos. A domain, on the other hand, is a collection of videos that all contain the same object type and are displayed as a collection.

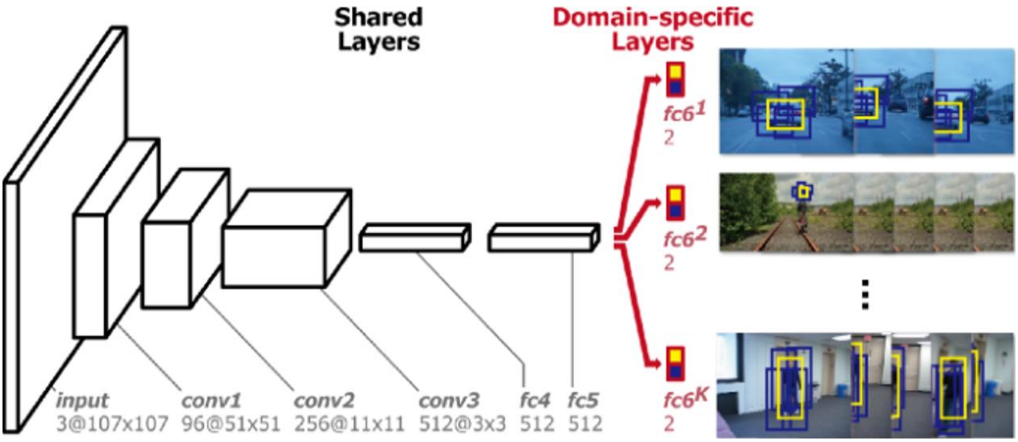


Figure 8: Branches of domain-specific layers

CLASSIFICATION-BASED SEGMENTATION

An observer matrix with each pixel acting as an observer is created in the previous step and used to generate the classification input in the following step. Following that, each observer is classified according to the variables that are present, using a learning model (such as the object and the background, or defects and defects, etc) (Du and Sun, 2006). When performing the classification process, it is common practice to provide a training set for the classification that contains images that have been successfully manually divided into the so-called supervisory learning. After obtaining the coefficients for the learning model, it is possible to classify the test picture set using the same model as before, but with the coefficients that were previously obtained. The supervised classification procedure is depicted in action in Figure 11, which illustrates the procedure. However, while the use of a training image is advantageous, it is not always necessary because unattended learning techniques such as clustering and the self-organization map are available that allow the observers to be classified into different classes without the need for any prior knowledge.

When using classification-based methods, one disadvantage is that it is necessary to know the goal of segmentation before segmentation can be carried out; in other words, it is necessary to know the number of classes into which pictures can be segmented before segmentation can be carried out.

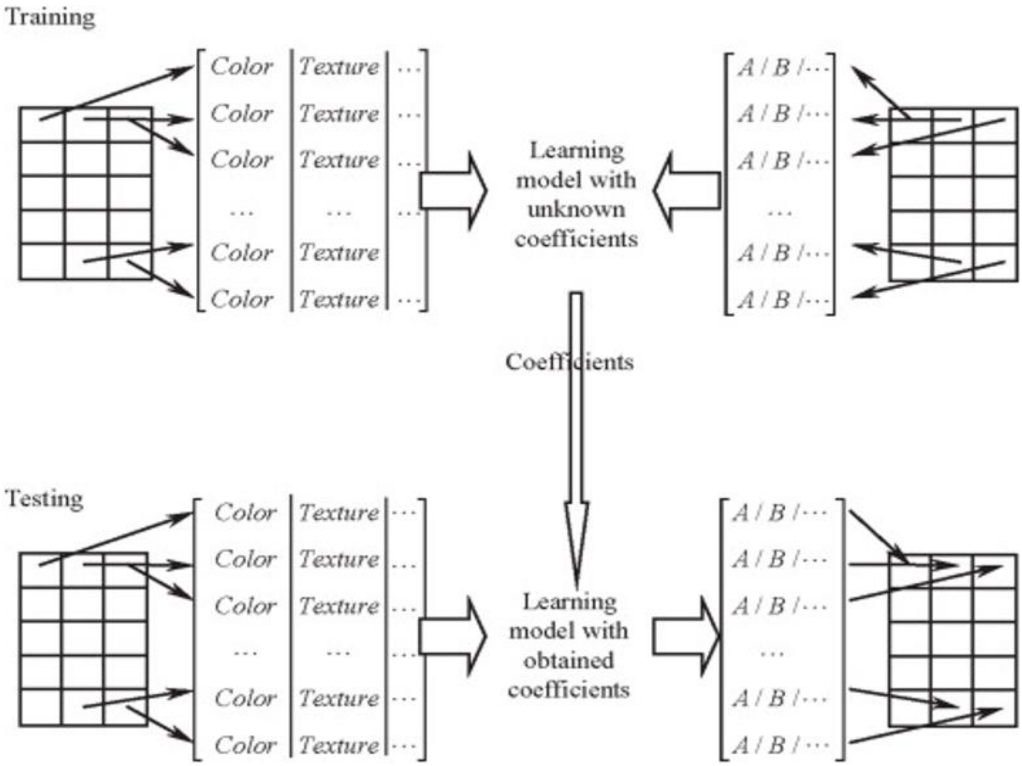


Figure 11: Classification-based segmentation

Image Extraction

Beyond being the most widely used and important classification characteristic, the value of pixel intensity also serves as the primary source of pixel information due to its high level of accuracy. When viewing a grayscale image, the intensity value for each pixel is one; when viewing a color image, the intensity value for each pixel is three; and when viewing a black and white image, the intensity value is one. Obtaining an intensity value from a single image using this technique, which allows for the acquisition of multiple images of the same product at different wavelengths, is becoming increasingly popular as an alternative to traditional methods. To determine the intensity of a product at different wavelengths, it is employed. Using the following procedure, it is possible to extract the intensity values of the same pixel as that of the pixel classification from a variety of images: Computer vision technology has piqued the interest of researchers who are interested in applying this technique to the task of determining the quality of apples (Leemans et al., 1999; Blasco et al., 2003; Kleynen et al., 2005). To obtain more information about the pixels, it may be necessary to extract the features of the pixels from only a small pixel-centered region, which may be possible depending on the situation. As a result, in addition to the intensity value, the image texture - which is a significant factor on the surface of the product for pattern recognition due to its powerful discrimination ability (Amadasun and King, 1989) - can be extracted as a classification feature for pixels in addition to the intensity value. Image texture is a significant factor on the surface of the product for pattern recognition due to its powerful discrimination ability. A significant factor on the surface of a product for pattern recognition is image texture, which is important due to its powerful discrimination ability on the surface of the product. Further technical information on the extraction of image texture features can be found in the review section of this document, which contains more in-depth technical information (Zheng et al., 2006).

Dimension reduction method

When there are a high number of data points in the classification input matrix, it is common practice to restore the matrix dimension to its original value before beginning the classification process. Principal component analysis (PCA) is a strong approach for assessing reduction that is most often used in classification research to minimize the number of classification variables that must be utilized in a particular study. To achieve classification segmentation, it is necessary to minimize the number of classification observers in use. PCA is not a good choice for classification and segmentation tasks. This is because the PCA does not allow for a reduction in the number of categorization observers to be used in the experiment. Researchers were able to create and put into effect the self-organizing map as a consequence of their findings (SOM). Using regularizations and correlations between observers, SOM may be expanded to produce an unregulated neural network, with each neuron representing a group of observers who have comparable variables. This results in a neural network that has neurons that are not under control. Using the SOM as a starting point and extracting the internal topological structure of the input matrix from regularization and correlation between observers, one can construct an uncontrolled neural network in which each neuron is composed of a group of observers with the same internal topological structure. When the SOM is used instead of the original observers, observers allocated to the neural class that they belong to can be used in place of the original observers, as seen in the following example (Chtioui et al., 2003; Marique et al., 2005; Bynagari & Fadziso, 2018).

Classification

So far, Bavaria's theory (the ST method) and fuzzy clustering (the ST and FL combination) are the only two approaches that have been proposed in the food industry, despite the fact that a variety of approaches, including statistical techniques (ST), neural networks (NN), vector support (SVM), and fuzzy logic, are now widely available (FL). It is possible to compute the Bayesian probability P(Ci|X) using the features (variety) of the following equation and the Bayesian theory, which indicates how likely it is that the next equation is related to the Ci-class pixels. The Bayesian probability P(Ci|X) is a probability that indicates how likely it is that the next equation is related to the Ci-class pixels (observer)

$$p(C_i|X) = \frac{P(X|C_i)P(C_i)}{P(X)}$$

An observer with variable X may belong to class Ci; nevertheless, it is not feasible that an observer with variable X belongs to class Ci. An observer with variable X may belong to class Ci. P(X|Ci) shows a prior possibility that an observer with variable X belongs to class Ci, and P(X|Ci) indicates a prior probability that an observer with variable X belongs to class Ci. P(X|Ci) indicates a prior possibility that an observer with variable X belongs to class Ci. An observation observer's prior likelihood that a variable X will be used by that observer before the observation happens is represented as the probability that the observation will occur (Ci). In the following phase, a Bayesian probability limit is calculated, and if the likelihood of an observer exceeding a specific threshold is exceeded, the observer is categorized as Class C, and vice versa, is determined (Neogy & Bynagari, 2018).

COMPUTER VISION APPLICATIONS IN 2021

Computer vision in sports

Player pose tracking: Machine learning techniques can be used to detect patterns between human body movements and pose in video or real-time video streams over a large number of frames using artificial intelligence vision. Using single stationary cameras above and below the water's surface, researchers have been able to estimate human poses, which they have then applied to real-world videos of swimmers, with encouraging results (Ganapathy et al., 2021a). Athletes' performance can be evaluated quantitatively using these videos because they do not require the manual annotation of body parts in each video frame, allowing for significant time savings. So convolutional neural networks are used to automatically detect information about the athlete's position as well as information about his or her swimming style, as a result.

Stroke recognition: Algorithms for computer vision can be used to detect and classify strokes in various applications (for example, classifying strokes in table tennis). It is necessary to perform additional interpretations of the identified instance and labeled predictions after the movements have been identified or classified (for example, differentiating tennis strokes as forehand or backhand). Aiming to provide teachers, coaches, and players with tools for analyzing table tennis games and improving sports skills, the Stroke Recognition system has been developed. It is currently being tested in beta.

Figure 12: Player Tracking With Deep Learning-Based Pose Estimation

Goal-line technology: To assist arbitrators in their decision-making, camera-based technologies can be utilized to identify whether or not an objective has been reached. When compared to sensors, the visual approach does not require any adjustments to the normal soccer equipment and does not require any additional training. They are based on high-speed cameras that triangulate the location of the ball on the goal line by comparing photos taken at different speeds. The development of an algorithm that scans applicant ball areas to recognize the pattern of balls is required to detect balls.

Computer Vision in Healthcare

Cancer detection: Machine learning plays an important role in the medical sciences like as skin and breast cancer detection. For example, scientist can detect minute changes in cancerous and non-cancerous images that help in the diagnosis of data by using MRI (magnetic resonance image) scans. And by analyzing it, benign or malignant images are inputted.

Covid-19 diagnosis: Corona virus control can also be ensured by using computer vision techniques. Multiple deep learning computer vision models exist for x-ray based COVID-19 diagnosis. Digital chest x-ray radiography images is one of the most common method named as COVID-Net. It was introduced by Darwin AI, Canada.

Tumor detection: Deep neural networks use in the detection of brain tumors. MRI scan makes it possible to detect. Tumor detection software is very helpful for doctors in detection of tumors at high accuracy. New techniques are being tested to elevate the accuracy of these diagnoses.

Computer Vision in Agriculture

Crop monitoring: Production and quality of crops (rice) determine the sustainability of food security. Subjective human judgment was the only method for crop monitoring before. It is inappropriate and time-consuming. Computer vision techniques ensure the determining of plant growth and nutrient requirements. Computer vision application spots the elusive changes in crops caused by malnutrition. It provides the accurate and timely regulation. It is used to measure the plant growth indicators and to regulate the stage of growth as well.

Insect detection: To control the pest, counting and recognition of flying insects ought to be done smoothly and immediately. It is inappropriate and time-taken method to count the insects manually. Vision based systems play an important role in identifying and counting of insects such as YOLO based system.

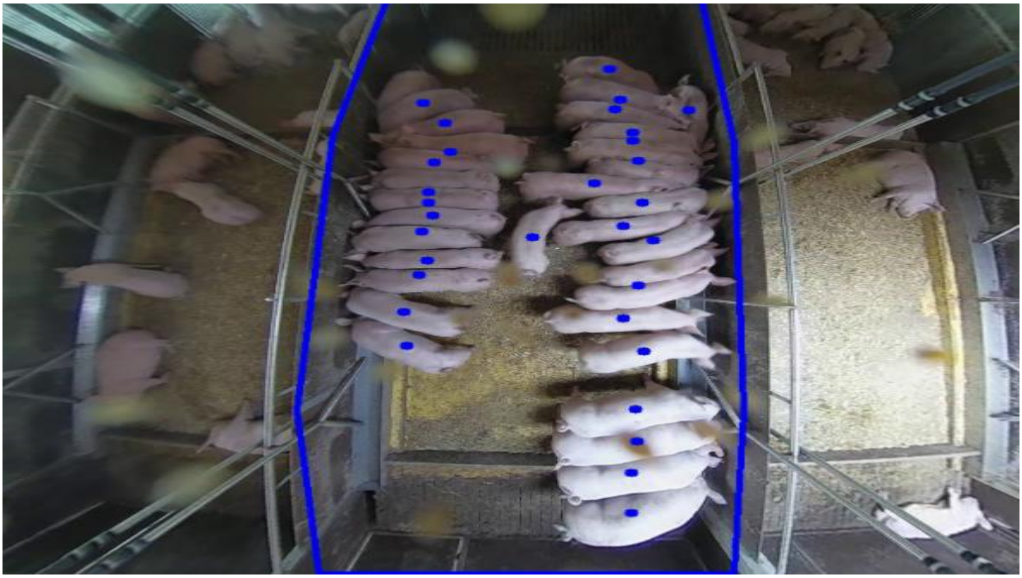


Figure 12: Agriculture Computer Vision Application for Animal Monitoring

Animal monitoring: Because of the computer vision technology, animal behavior and monitoring can be checked on daily-basis. Birthing, illness detection and behavioral changes are monitored remotely by using this cutting-edge technology. Without coming close to or touching the animal, scientist can examine the wildlife animals from a safe distance.

Computer vision in transportation

Vehicle classification: Vehicles can be detected, tracked, and classified in a variety of ways. Computer vision-based-sensors like as closed circuit television cameras (CCTV), LiDAR (light detection and ranging) and thermal imaging devices. By joining all these systems in a vehicle can increase the categorization accuracy of vehicle. In addition to these systems, various computer vision-based technologies in construction truck recognition are being used. Productivity evaluation, safety monitoring and managerial decision-making have been made possible by using these techniques (Ganapathy et al., 2021b).

Traffic flow analysis: Computer vision-based techniques like as intrusive and non-ivasive technologies have been reconnoitered in ITS (intelligent transportation systems). As the computer vision and AI innovates, video analytics are being used in omnipresent traffic cameras. This has great impact on ITS and smart cities.

Driver attentiveness detection: Distracted driving — such as daydreaming, talking on the phone, or glancing out the window – is responsible for a significant number of road traffic fatalities around the world. Artificial intelligence is being utilized to better understand driving habits and develop solutions to reduce traffic accidents. Passenger compartment violations are monitored using road surveillance technologies (Bynagari, 2019).

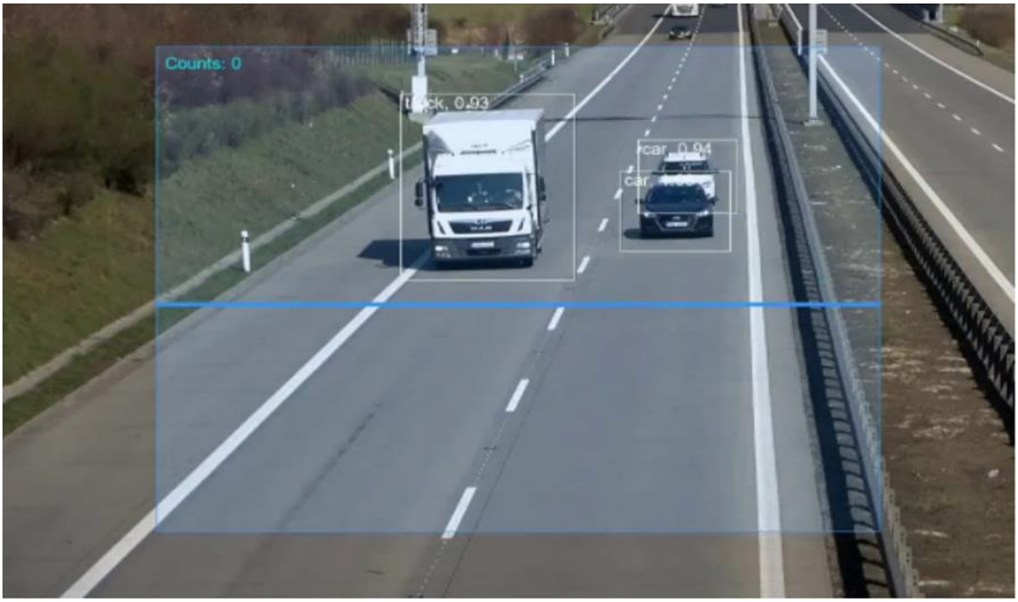


Figure 13: Computer Vision Application for Vehicle Counting

Conclusion

Based on the foregoing explanation, it is obvious that computer vision technology will continue to be a valuable tool for addressing a wide range of meat classification and quality prediction concerns. In stories on new technology, computer vision is a prominent topic. This technology differs from others in that it takes a distinct approach to data. Massive amounts of data that we generate on a daily basis, which some perceive as our generation's curse, are also used to our advantage: the data can educate computers to see and understand objects. This technique also represents a significant stride forward in our civilization's development of artificial intelligence. For additional analysis, image segmentation is used to separate food components from the backdrop. Threshold-based segmentation divides images into histograms using an optimal threshold determined by manual selection, the isodata algorithm, objective functions, clustering, and a variety of other methods. To fix segmentation errors, image-closing and -opening are occasionally used.

REFERENCES

Ahmed, A. A. A.; Paruchuri, H.; Vadlamudi, S.; & Ganapathy, A. (2021). Cryptography in Financial Markets: Potential Channels for Future Financial Stability. Academy of Accounting and Financial Studies Journal, 25(4), 1–9. https://doi.org/10.5281/zenodo.4774829

Amadasun, M., & King, R. (1989). Textural features corresponding to textural properties. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(5), 1264–1274

Bakirtzis, A., and Spyros, K. (2016). Genetic algorithms. Advanced Solutions in Power Systems: HVDC, FACTS, and Artificial Intelligence: HVDC, FACTS, and Artificial Intelligence, 845-902. DOI: https://doi.org/10.1002/9781119175391

Blasco, J., Aleixos, N., Moltó, E. (2003). Machine vision system for automatic quality grading of fruit. *Biosystems Engineering*, *85*(4), 415–423.

Bynagari , N. B. (2020). The Difficulty of Learning Long-Term Dependencies with Gradient Flow in Recurrent Nets. *Engineering International*, *8*(2), 127-138. https://doi.org/10.18034/ei.v8i2.570

Bynagari, N. B. (2018). On the ChEMBL Platform, a Large-scale Evaluation of Machine Learning Algorithms for Drug Target Prediction. *Asian Journal of Applied Science and Engineering*, *7*, 53–64. Retrieved from https://upright.pub/index.php/ajase/article/view/31

Bynagari, N. B. (2019). GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. *Asian Journal of Applied Science and Engineering*, *8*, 25–34. Retrieved from https://upright.pub/index.php/ajase/article/view/32

Bynagari, N. B., & Amin, R. (2019). Information Acquisition Driven by Reinforcement in Non-Deterministic Environments. *American Journal of Trade and Policy*, *6*(3), 107-112. https://doi.org/10.18034/ajtp.v6i3.569

Bynagari, N. B., & Fadziso, T. (2018). Theoretical Approaches of Machine Learning to Schizophrenia. *Engineering International*, *6*(2), 155-168. https://doi.org/10.18034/ei.v6i2.568

Chtioui, Y., Panigrahi, S., Backer, L. F. (2003). Self-organizing map combined with a fuzzy clustering for color image segmentation. *Transactions of the ASAE*, *46*(3), 831–838.

Du, C-J, & Sun D-W. (2006). Automatic measurement of pores and porosity in pork ham and their correlations with processing time, water content and texture. *Meat Science, 72*(2), 294–302.

Ganapathy, A., Hossain, M. S., Rahman, M. M., Asadullah, ABM., Amin, R. (2021b). The Significant of Biases in Learning Algorithms Generalization. *International Journal of Aquatic Science*, *12*(2), 3042-3052. http://www.journal-aquaticscience.com/article_134827.html

Ganapathy, A., Vadlamudi, S., Ahmed, A. A. A., Hossain, M. S., Islam, M. A. (2021a). HTML Content and Cascading Tree Sheets: Overview of Improving Web Content Visualization. *Turkish Online Journal of Qualitative Inquiry*, *12*(3), 2428-2438. https://www.tojqi.net/index.php/journal/article/view/1724

Khan, W., Ahmed, A. A. A., Vadlamudi, S., Paruchuri, H., Ganapathy, A. (2021). Machine Moderators in Content Management System Details: Essentials for IoT Entrepreneurs. Academy of Entrepreneurship Journal, 27(3), 1-11. https://doi.org/10.5281/zenodo.4972587

Kleynen, O., Leemans, V., Destain, M-F. (2005). Development of a multi-spectral vision system for the detection of defects on apples. *Journal of Food Engineering*, *69*(1), 41–49.

Leemans, V., Magein, H., & Destein, M-F. (1999). Defect segmentation on 'Jonagold' apples using color vision and a Bayesian classification method. *Computers and Electronics in Agriculture*, *23*(1), 43–53.

Marique, T., Pennincx, S., Kharoubi, A. (2005). Image segmentation and bruise identification on potatoes using a Kohonen's self-organizing map. *Journal of Food Science*, *70*(7), E415–E417.

Matiacevich, S., Celis Cofré, D., Silva, P., Enrione, J., Osorio, F. (2013). Quality Parameters of Six Cultivars of Blueberry Using Computer Vision. International Journal of Food Science. https://doi.org/10.1155/2013/419535

Nandakumar, V., Hansen, N., Glenn, H. L., Han, J. H., Helland, S., Hernandez, K., Senechal, P., Johnson, R. H., Bussey, K. J. and Meldrum, D. R. (2016). Vorinostat differentially alters 3D nuclear structure of cancer and non-cancerous esophageal cells. Scientific reports 6. https://doi.org/10.1038/srep30593

Neogy, T. K., & Bynagari, N. B. (2018). Gradient Descent is a Technique for Learning to Learn. *Asian Journal of Humanity, Art and Literature*, *5*(2), 145-156. https://doi.org/10.18034/ajhal.v5i2.578

Patel, K. K., Kar, A., Jha, S. N., and Khan, M. A. (2012). Machine vision system: a tool for quality inspection of food and agricultural products. Journal of food science and technology, 49(2), 123-141. https://doi.org/10.1007/s13197-011-0321-4

Zheng, C., Sun, D-W., Zheng, L. (2006). Recent applications of image texture for evaluation of food qualities – a review. *Trends in Food Science & Technology*, *17*(3), 113–128.

--0--

Archive Link: https://abc.us.org/ojs/index.php/ei/issue/archive