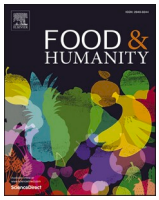
Food and Humanity 3 (2024) 100365



Contents lists available at ScienceDirect

Food and Humanity

journal homepage: www.editorialmanager.com/foohum/journal_overview.html



# Deep learning techniques for the exploration of hyperspectral imagery potentials in food and agricultural products

Abdulwahab Ismail Durojaiye [a,*], Samuel Tunde Olorunsogo [b], Bolanle Adenike Adejumo [b], Alkali Babawuya [c], Ida Idayu Muhamad [d]

[a] Department of Agricultural and Bioresource Engineering, Faculty of Engineering and Engineering Technology, Abubakar Tafawa Balewa University, P.M.B. 0248, Yelwa Campus, Bauchi State, Nigeria
[b] Department of Agricultural and Bioresource Engineering, School of Infrastructure, Process, Engineering and Technology, Federal University of Technology Minna, Nigeria
[c] Department of Mechatronics Engineering, School of Infrastructure, Process, Engineering and Technology, Federal University of Technology Minna, Nigeria
[d] Department of Bioprocess Engineering, Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, Malaysia

ARTICLE INFO

*Keywords:*
Chemometric and morphometric properties
Optical device
Spectral parameter
Hyperspectral imaging (HSI) system
Machine vision system

ABSTRACT

Global attention on the exploration of hyperspectral imaging (HSI) system for non-destructive evaluation of detailed analyses of food and agricultural products has increased significantly. This is due to its feasibility of obtaining spectral signatures (SS) of various substances. HSI is a special type of camera capable of simultaneously capturing high precision chemometrics features for the recognition of objects pixel by pixel. Unlike the digital camera equipped with the ability to obtain the morphometric properties of an object in term of RGB (red, green, and blue) colours. High-cost implication of HSI configurations mostly hindered researchers, institutes and professionals' accessibility to this promising optical equipment. Hence this paper provides different techniques commonly employed for developing low-cost commercial off-the-shelf (COTs) and do-it-yourself (DIY) embedded artificial intelligence HSI system for future exploration potentials of handling numerous food and agricultural materials. Relevant multivariate algorithms for ease of HSI utilisation are extensively highlighted to ensure seamless future adoption. This survey, to the best of our knowledge, addressed requirements necessary for ease of HSI replication in any part of the world. It suggest possible modifications necessary for HSI development and articulates numerous techniques adopted in exploring hyperspectral imagery as an artificial intelligence technology for handling various food and agricultural materials as an embedded system.

## 1. Introduction

Hyperspectral image (HSI) camera is generally made up of an objective lens (either C–mount, F – mount, or S–mount type), an imaging spectrograph (input slit + collimator lens + diffraction grating lens), a detector (CCD-charge couple device or CMOS-complementary metal oxide semiconductor) and a focus lens (Fig. 1). The HSI configuration mostly possessed the potential of obtaining high- precision detection of spectral and chemometric analysis for recognizing features of objects pixel by pixel unlike the digital camera which consists mainly of the objective lens and the grayscale optical device. HSI complete system can be embedded with a sample translation or rotational stage, illumination source, mounting tower, and computer-supported image acquisition software (Fig. 2). Researchers opined that the construction of hyperspectral imagery concerning the cost implications depends on the quality, availability, affordability, and selection of commercial off-the-shelf (COTs) optical components for specific targets (indoor or outdoor purposes) (Pechlivani et al., 2023a). Others revealed that the quality of hyperspectral images is mostly determined through the detector's performance. Universally, a highly sensitive detector with elevated signal-to-noise ratio (SNR) which enhances the image preprocessing power to extract 3D (spatial and spectral parameters) information from the acquired data is frequently preferred (Saha & Manickavasagan, 2021). The unique choice of design by researchers usually contributes to the variations in the hyperspectral device configuration with respect to the total cost implication. Therefore increase demand for hyperspectral imagery across the globe has attracts and encouraged researchers efforts to continuously developing the

* Corresponding author.
*E-mail address:* adismail@atbu.edu.ng (A.I. Durojaiye).

https://doi.org/10.1016/j.foohum.2024.100365
Received 20 May 2024; Received in revised form 22 July 2024; Accepted 23 July 2024
Available online 25 July 2024
2949-8244/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

emerging technology for efficient utilisation in various field of specialisation. The smart system has demonstrated it rapidity and accuracy in various applications and has call the attention of end users across the world. Hence, continuous production of HSI will further increase wider awareness on the exploration potentials of the trending technology and ensure precision in agriculture, healthcare, food safety, environmental monitoring, defense and applications in food industries. Therefore the need for modification of HSI system is required where the central processing unit (CPU), the monitor and the image acquisition component is integrated together as an embedded system to enhance portability and computational flexibility.

However, despite numerous feature exploration potentials and advantages of HSI system over other conventional imaging systems, the high cost of acquiring HSI can be exorbitant. The commercially available HSI system range between USD 10,000 to USD 100,000 (Grigoriev, 2022). High cost of the HSI has greatly encouraged researchers' ability to exhaustively and consistently intensified efforts in obtaining cheaper and COTs material components for integration and development of low-cost HSI system suitable for various applications. The schematic representation of the core configuration of a push broom HSI system is shown in Fig. 2.

Research revealed that HS imagery depending on the specific area of applications can possess varying wavelength ranges from the visible spectrum (380 – 740 nm), while some others extend into near–infrared (NIR, 700 – 2500 nm), or short–wave infrared (SWIR, 1000 – 2500 nm) regions. Other imagery devices can cover the thermal infrared (TIR, 8000 – 15,000 nm) region. For visible and near–infrared (NIR) wavebands, HSI can have a wide field of view (FOV) usually varying from few degrees to over 100 degrees (Feng et al., 2019). The HSI device mostly operates within spectral regions of ultraviolet (UV, 200 – 400 nm), visible and near-infrared (VNIR, 400 – 1000 nm), and near-infrared (NIR, 900 – 2500 nm) but are commonly active at visible and near-infrared spectrum (Fig. 3). HSI specifically, can be designed to satisfy various targets depending on the researcher's area of application. Therefore, most pushbroon HSI system are commonly configured to operates with the spectral range of visible to near infrared region for their respective effectiveness depending on the food and agricultural materials brought under evaluation. Thus, the need for modification of HSI system is required where the graphical user interface (GUI), the central processing unit (CPU) and the image acquisition component is integrated together as an embedded system to enhance portability and computational flexibility.

The majority of the researchers aligned with the pattern of arrangement of the selected HSI materials as illustrated in Fig. 4. The orientation includes the area of energy source (light) ($L_s$), half angles ($\alpha$), objective lens (front lens) ($L_o$), collimator lens or mirror ($L_c$) and focusing lens or mirror ($L_f$). The organisation also comprises of aperture stop ($A_s$), entrance penetration area of the source image ($P_a$), area of entrance slit ($E_s$), area of the diffracted slit image (exit slit area) ($E_e$), diffraction grating prism ($D_g$), objective distance ($O_d$), image distance ($I_d$) of $L_o$, focal length ($F_l$) of $L_c$ and focal length ($F_l$) of $L_f$.

The HSI optical device allows standards wavelength acquisition by dispersion of light through a diffraction grating ($D_g$) (Eismann, 2012). Light is observed to often enters a series of lenses (an objective) $L_0$ that must focus the light source to a very thin slit. The slit further permits only a thin line of light to pass through to the next objective lens ($L_c$). This objective collimates the slit of light that reaches the grating($D_g$). The grating then separates the light rays just like a prism, and the separate wavelengths pass through a final objective lens ($L_f$) at a refraction angle ($\alpha_3$) based on the specification of the objective lens selected. This objective directs the image light rays onto a CMOS image sensor detector (D) through $E_e$, where the photons are converted to digital units. The incorporated internal software then handles the 3D acquired signal by the optical instrument for further processing in preparation for downlink. Research revealed that grating efficiency and the quantum efficiency of the detector account for the spectrum parameters of the target objects. Likewise, the greater the slit width, the more the light received by the system which positively affects the signal-to-noise-ratio (SNR) and consequently reduces spatial resolution. The result of the 3D signal (image) received from the transmission grating instrument (a spectrogram) consists of unique spatial and spectral dimensions. The spectrogram reveals the actual image combined with the adjacent strips (with additional spectral information useful for identifying the target object). Sigernes et al. (2018) attested that the basic optical components for HSI development consist of a front lens $L_0$, one high-precision slit, and a collimator lens ($L_c$). The combination necessitates the camera to require three thin lens holders, one focus tube, one lock ring, and a detector lens. Fig. 4 summarises the global acceptable orientation of an hyperspectral imagery pattern of arrangement to ensure deep ray penetration and analysis for proper interpretation and visualisation. Therefore, this reviewed paper highlights the exploration potentials of deep learning techniques mostly adopted for hyperspectral imagery analyses of food and agricultural products.

Over some decades, the conventional means of evaluating the nutritional constituents of food and other agricultural products are widely used in industries and research institutes. This evolves the addition of chemical reagents and subjecting the sample to hydrothermal treatment and hence manually calculating the nutrientional parameters of the food samples. The methods are time-consuming, tedious, prone to error and are called destructive methods. Currently,

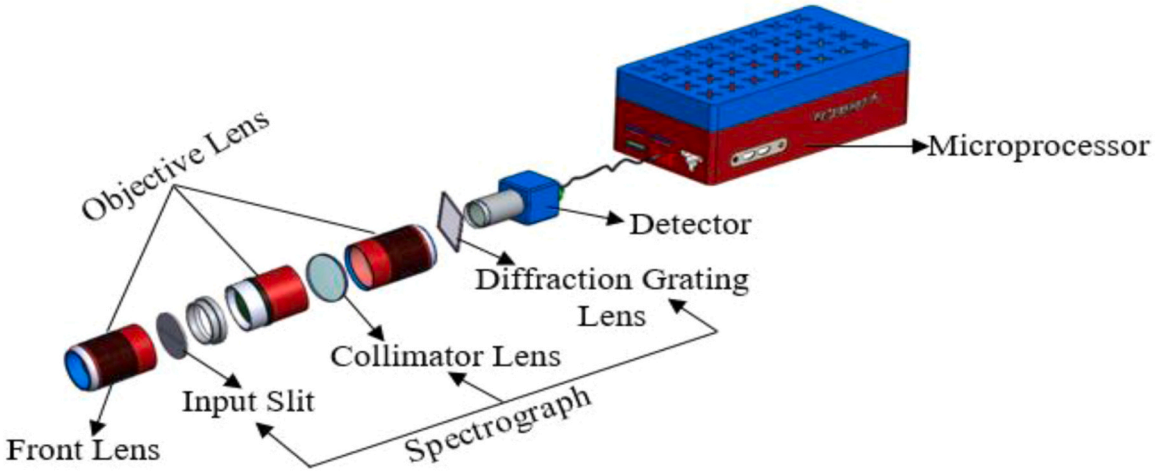



**Fig. 1.** Exploded view of HSI camera architecture indicating the optical components.

technological advancement has greatly assisted researchers leading to the development of numerous machine vision and artificial intelligence systems, precision agriculture, remote sensing imagery (RSI), and smart farming tools for handling different food and agricultural materials (Henriksen et al., 2022). Many artificial intelligence (AI) tools have emerged. Some of which are suitable for application in various fields of specialisation ranging from food and agric-products analyses, aeronautics, surveying, transportation, environmental sanitation, health care, defense and communication. These equipment are referred to as non-destructive tools as a product of AI configurations. Among known emerging image–based techniques suitable in the food processing field are hyperspectral imaging (HSI) smart system, laser-Induced breakdown spectroscopy (LIBS), laser backscattering image (LBI), laser-light backscattering imaging (LLBI) (Kabir et al., 2022). Other image technology for obtaining food quality analyses and safety also include stereo systems (SS), thermal infrared (TIR), remote sensing image (RS), thermal or traditional imaging (TI), X–ray imaging (XRI), fluorescence imaging (FI), magnetic resonance imaging (MRI), Raman imaging (RI), ultrasound imaging (UI), odour imaging (OI) and microwave imaging (MI) (Zhu et al., 2021; Hussain et al., 2018). Among all the technologies, HSI system has gain wider area of applications which include precision agriculture, food quality control, disease detection in fruits and vegetables, environmental monitoring, and biomedical imaging examinations (X-ray) (Pechlivani et al., 2023). HSI has also displayed unique potentials in the area of soil property analyses (soil nutrient availability, organic matter content and moisture level). This information will greatly improve fertilizer applications to crops and enhance soil productivity thereby optimising crop yields, provided adequate exploration is conducted by the farmers to ensure food security (Thomas et al., 2018). Othman et al. (2023) revealed that convolutional neural network (CNN), deep neural network (DNN), recurrent neural network (RNN), support vector machine (SVM), artificial neural network (ANN), random forest (RF) and several others being an artificial intelligence models has significantly contributed to food and agricultural products quality detection and adulteration. Therefore, AI ideas has greatly assist researchers in the food and postharvest technology realms to developed meaningful technological tools between 2015 through 2022 and still trending.

## 2. Selection of quality detector for HSI development

Several image-detecting sensors are commercially available for the construction of hyperspectral imaging cameras from imaging development systems (IDS), all of which perform similar functions. Some available detectors include but are not limited to Sony IMX174CMOS image sensor with high quantum efficiency (QE), Sony IMX249, Sony IMX252, and CMOSIS CMV2000 sensors (Henriksen et al., 2022). Others are Sony Super HAD colour CCD, 5 M pixel colour CMOS, and Monochrome CMOS (Sigernes et al., 2018). Research revealed that the Sony IMX249 is a less expensive version of the IMX174 with similar optical performance but possessed low frame rate. The Sony IMX252 and CMV2000 are categorically more sensitive in the NIR wavelengths region at the cost of having lower quantum efficiency (QE) in the VIS range. Other parameters required to be considered during sensor selection include pixel size, pixel pitch, sizes of the sensor, number of pixels, well depth, and bit depth among others. Another important factor contributing to the image acquisition quality is the magnification lens (COTs available ones are 10 ×, 40 ×, 70 × and 100 ×), aperture, and focal length settings (ranging from F/1.7 to F/16). Pechlivani et al. (2023), selected a 10 MP (Megapixel) resolution objective lens (10 × magnification telephoto) with a focal length of 32 mm with an adjustable aperture ring to allow the front objective lens provides capabilities, precise and accurate focus distant objects. The researchers utilised a Raspberry Pi NoIR camera sensor equipped with an 8 MP Sony IMX219 CMOS Sensor which is capable of offering an image area of 3.68 10 × 2.76 mm and pixel dimensions of 1.12 μm 10 × 1.12 μm. Due to large data acquisition and storage capacity, image processing and analysis are required by HSI. Consequently, Raspberry Pi 4 microprocessor (with 8 GB RAM, and 16 GB internal storage with MicroSD) was selected by the research team to enhance the processing speed and rapid output unlike other microcontrollers suitable for tiny machine learning processing. The Raspberry Pi 4 used was known as a single-board computer (SBC) or mini-computer due to its Broadcom BCM2711, quad-core Cortex-A72 (ARMv8) 64-bit SoC @ 1.5 GHz processor, it is suitable and adequate for tasks that required significant computing power, as well for handling multiple live data in the computer vision system, deep learning and machine learning projects (Norris, 2020). It is however, evident that machine vision system has been widely adopted by researchers across the globe for handling complex analysis for ease of operation both on the factory floor and for research purposes which ordinarily are tedious and cumbersome to execute manually.

### 2.1. Variety of filaments for HSI 3D printer

After careful selection and arrangement of optical components of the HSI camera, the device is usually secured with a casing in a 3D printer using various types of filaments. The exterior sheath commonly used for securing the cameras by most researchers are made up of three major filaments namely: polyethylene terephthalate glycol (PETG), polylactic acid (PLA), and anti-bacterial sheet scientifically called acrylonitrile

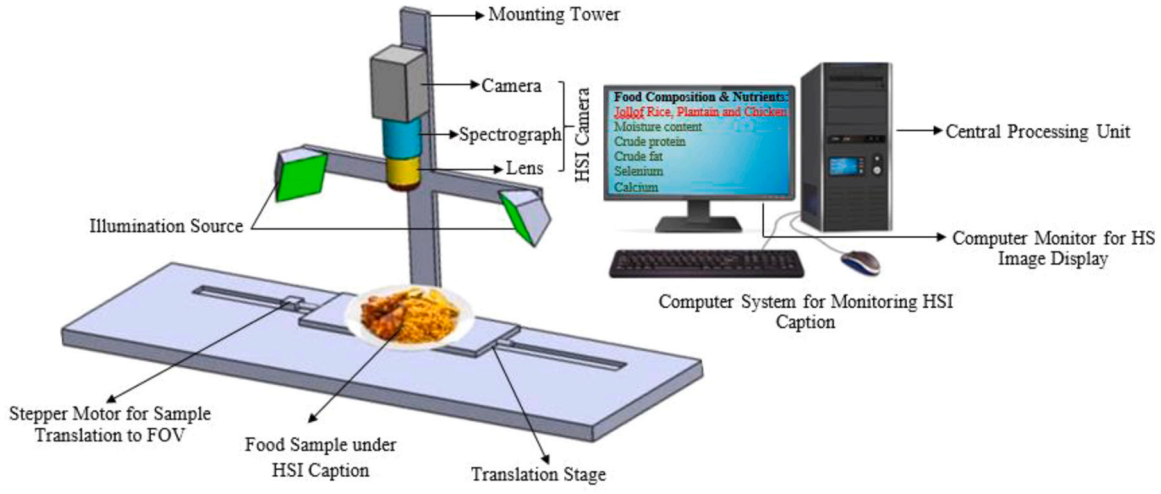



**Fig. 2.** Schematic picture of the HSI system used for food sample assessement: FOV- field of view, HSI – hyperspectral imaging.

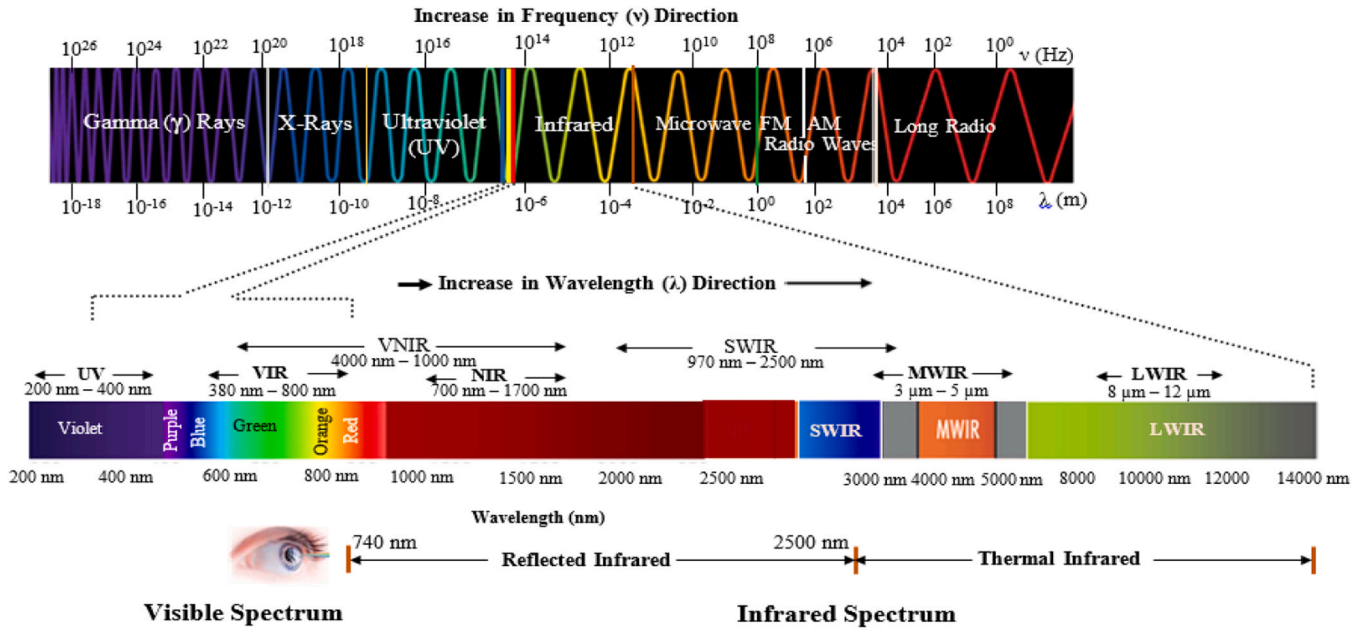



**Fig. 3.** Electromagnetic spectrum wavelengths for nondestructive assessments: UV – ultraviolet, VIR – visible infrared, VNIR – visible and near infrared, NIR - near infrared, SWIR – shortwave infrared, MWIR – mid-wave infrared, LWIR – longwave infrared, SWIR – shortwave infrared, nm – nanometer, μm – micrometer, v-velocity, Hz – hertz.

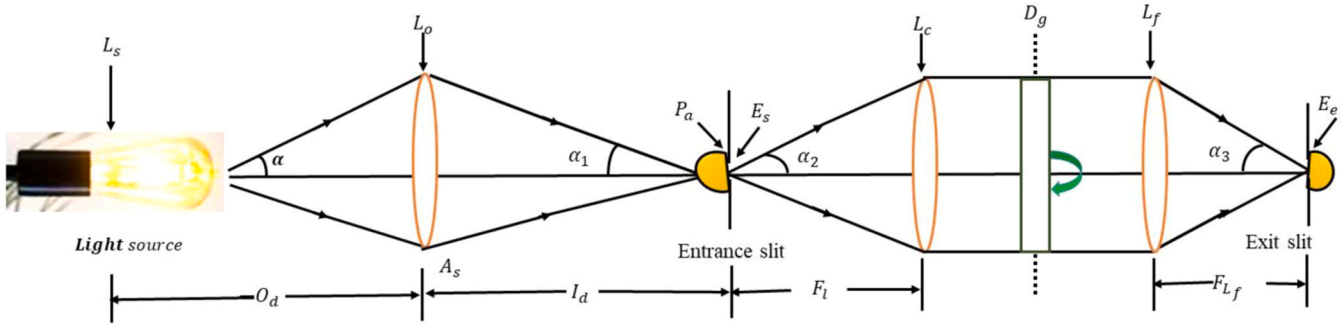



**Fig. 4.** Arrangement of hyperspectral imaging camera optical components: $L_s$ – light source, $L_o$ – front objective lens, $P_a$ – light penetration area, $E_s$ – entrance slit area, $L_c$ – collimator lens, $D_g$ – diffraction grating prism, $L_f$ – final objective lens, $E_e$ – exit slit area, $O_d$ – objective distance, $l_d$ – image distance, $F_t$ – focal length, $F_{lf}$ – focal length of diffraction prism, $\alpha$ – half angle, $\alpha_1$ – entrance slit angle, $\alpha_2$ – exit slit angle, $\alpha_3$ – refraction angle of an objective lens.

butadiene styrene (ABS). ABS (contains 20 % acrylonitrile, 25 % butadiene, and 55 % styrene) and PETG has been confirmed for heat resistance, waterproof, durable, reusable, recyclable, suitable, and safe for use in food process due to the material compositions rendering the filaments compatible for use in high moisture and intense sunlight environment (Hsueh et al., 2021; Pechlivani et al., 2023b; Petrov et al., 2021). Meanwhile, the image quality to be obtained from the HSI system strongly depends on a fixed optical train (Høye et al., 2015; Skauli, 2017). (Prentice et al., 2021) opined that there is difficulty identifying any unknown effects of component shift, vibration, or rotation after the imager has been secured by the 3D printer. Therefore adequate care is encouraged during the printing process to avoid extreme vibrations and shock during operation, it is therefore essential to limit the complexity, and number of parts to bind along the train to ensure ease in fixing the components by the printer while avoiding distortion that could rattle loose in the sequence of event. Mao et al. (2022) for instance advanced an OpenHSI, designed to make compact push-broom HSI spectrometers more easily accessible to a broader research world by making available an open-source optical design and software platform. The researchers adopted the usage of two printing methods, where fused deposition of ABS was initially used and later used an SLA printer photosensitive resin. Fused deposition modeling (FDM) also known as fused filament fabrication (FFF) and stereolithography (SLA) among the major and popular types of 3D printers readily available in the commercial market. Both 3D printing technologies have been widely adopted and refined for the desktop, are easy to use, and are affordable in price. The device covers a spectral range of 400 nm – 830 nm with a total weight of 119 g. (Sigernes et al., 2018), in a research study demonstrated the design of a push–broom HSI camera using commercial off–the–shelf components with low-cost objective lens (camera head) and a 3D printed mount. The total cost of the device amounts to USD 700 or EUR 647.84 and the total weight is 200 g. The spectral wavelength ranged from 400 nm – 800 nm and was printed with PLA filament. Henriksen et al. (2022) also constructed a DIY visible–near-infrared (VIS/NIR) push broom HSI imager with C –mounted optics useful for dark surfaces especially deep ocean detection. The device has a total mass of about 650 g which cost USD 2600 with a spectral range of 400 nm – 800 nm although the filament type used for the 3D printing was not disclosed in the report. Summary of do-it-yourself hyperspectral imaging devices designed by distinct research teams and the materials used with respective cost implications are illustrated in Table 1.

**Table 1**
Various hyperspectral imagery designs with to spectral specifications and filament types.

| S/N | Authors | Total Cost (USD / EUR) | 3D filament Used | Total weight (g) | Spectral Resolution Range (nm) | Spatial Resolution (Pixels) | Size (cm) |
|---|---|---|---|---|---|---|---|
| 1 | Pechlivani et al. (2023)a | (194.56 EUR) 210.15 USD | PETG & PLA | 1263.65 | 379 – 937 | 127 × 125 | |
| 2 | Henriksen et al. (2022) | | | | 400 - 800 | | |
| 3 | Mao et al. (2022) | | ABS/ Photosensitive resin | 119 | 400 - 830 | 2064 × 1544 | |
| 4 | Riihiaho et al. (2021) | 2000 EUR 2160.29 USD | PLA | | 400 - 600 | 3376 × 2704 | |
| 5 | Botero-Valencia and Valencia-Aguirre (2021), | 251.83 EUR 272.01 USD | ABS | | 410 - 940 | | |
| 6 | Prentice et al. (2021) | | | | 300 – 1000 | 1936 × 1216 | |
| 7 | Salazar-Vazquez & Mendez-Vazquez (2020) | 500 USD | PLA | 300 | 400 - 1052 | 116 × 110 | |
| 8 | Sigernes et al. (2018) | 626 USD | PLA | 200 - 306 | 400 – 800 | | 15 × 6 × 5 |
| 9 | Uto et al. (2016) | 10,000 (USD) | | | 340 –820 | 18 × 1 | |
| 10 | Uto et al. (2015) | 1000(USD) | | 1200 | 340 - 750 | | |
| 11 | RalfHabel et al. (2012) | 2000(USD) | | | 410 - 700 | 120 × 120 | 15 × 12 × 25 |

## 3. Hyperspectral imaging (HSI) compatible software

Constant and rapid software package development has greatly assisted researchers in the selection of powerful toolboxes capable of handling image processing. Most HSI analyses are conducted using Environment for Visualisation Images (ENVI), Exelis Visual Information Solutions, Boulder, CO, USA is one of the influential tools for the analysis of hyperspectral and multispectral images, which can provide many functions including data transformation, filter, classification, visualisation and many other image manipulations. Independent Development Environments (IDE's) of the Python programming language platform can also be explored for image pre-processing, feature extraction, visualisation and model prediction. Mathematical laboratory – MATLAB (The Math Works, Inc., Massachusetts, USA), which is also a powerful image processing tool, can help to develop image processing routines using high–level technical computing language and interaction environment for data analysis, developing models and visualisation of hyperspectral image data. Decipher (CAMO Software AS, Oslo, Norway) is another chemometrics tool, it gives users the power of multivariate data and interactive data visualisation in an easy–to–use programme. Other useful packages which can greatly assist researchers in handling HS imagery are MIA_Toolbox (Eigenvector Research, Wenatchee, WA, USA), which analyses hyperspectral images using familiar PLS_Toolbox tools (http://www.eigenvector.com/software/pls_toolbox.htm); and hyperSpec (R package) to handle hyperspectral images (https://cran.r-project.org/web/packages/hyperSpec/index.html). Any one of the identified software must be deployed on the processor to enable the functionality of HSI systems. Once the system is fully designed and installation of the software is conducted, then the artificial intelligent smart system become efficient based of the technical knowledge of image processing through deep learning operation and machine learnig techniques.

### 3.1. Calibration of hyperspectral imaging (HSI) camera

All HSI developers acknowledged that once the device is fully designed, spectral images can be acquired as the system is turned on and most importantly, the device must be calibrated. Mostly square HSI software are commonly used for the operation. During the calibration process, the spectral images normally consists of two distinct regions namely: the zero-order mode (ZM) and the first-order mode (FM) regions (Habel et al., 2012). The ZM area refers to instances when light is not diffracted and appears as a small square located at the centre, while the FM or region of interest (ROI) is the relevant first time light is diffracted in the system (Fig. 5). It has higher intensity and its perimeter contains the ZM parameters and adequate amount of information which required energy at a specific wavelength. The identification of the ZM, ROI (FM), and other relevant references empowers the Client (squareHSI) software to estimates the parameters required to export HSI data. The process is known as calibration. Unlike other digital cameras, this calibration involves spectral properties of the diffraction system of any HSI configuration. It usually does not consider the variation of light which is often alleviated using normalisation based on white and dark references (ElnMasry et al., 2007).

During calibration, three spectral properties (SPs) are required to identify the location of spectra references. The parameters of the ROI and ZM have to be identifiable in the first SP (as indicated in Fig. 5a). The second SP must also contain a diffracted light with three or more well-known wavelength peaks. (Salazar-Vazquez & Mendez-Vazquez, 2020) uses light of a fluorescent lamp *Sylvania Delux EL* 2700 K spotlight (Fig. 5b) because it possesses five identifiable wavelength peaks at 405 nm, 440 nm, 545 nm, 615 nm, and 710 nm. Another author revealed that fluorescent lamps provide recognisable wavelengths within the desired range of 400 – 1052 nm and an incandescent lamps which emit a continuous spectrum (Abdel-Rahman et al., 2017). Then the third SP must contain a diffracted light with a continuous spectrum.

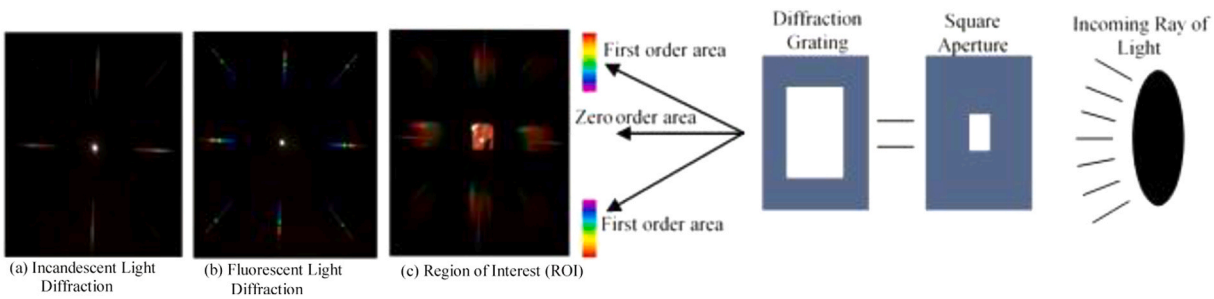



**Fig. 5.** Zero and First order (ROI) diffraction modes of the HSI system.

The researchers also used an incandescent lamp because it contains a light spectrum with energy in the whole operation range Raspicam (Fig. 5c). Additionally, the origin of both light sources (incandescent and fluorescent lamp) must be identified to distinguish the respective centre with the coordinate of the light origin and coordinate of the diffracted wavelength peaks. However, the squareHSI software is capable to estimate the relationship between wavelength – to – pixels and vice versa. Subsequently, the Client software utilised the relationship and continuous spectrum of light (Fig. 5c) to calculate the independent performance of the red, green and blue sensors of the imaging sensor inside the HSI device. Although Pechlivani et al. (2023) use 6500 K fluorescent lamps and concluded that either 3000 K incandescent lamps or 3000 K fluorescent lamp light source will also provide adequate spectral information required for the calibration. Therefore, calibration of HSI can be seamlessly conducted provided that the spectral parameter and wavelength range of the energy source used are identifiable which is essential for the device sensitivity within the VIS to NIR region.

### 3.2. White and black calibration on HSI device

Pixel–level hyperspectral image calibration otherwise referred to as white and black calibration is mostly conducted on images acquired under the laboratory scale. It enables the removal of random noise signals (dark current) and also ensure the conversion of acquired light intensity into standard reflectance/transmittance value. The dark current is caused by the light source or power supply and can be generated due to the heat of internal CCD or CMOS sensor and even in the absence of light (Wu et al., 2012). The other reason for the removal of a dark background is because the pixel unit retains harmful dark current and adds it to the produced signal. Hence, to correct the influence of such dark current, the white and black reference images are to be captured independently from an extreme illumination situation. Therefore, a standard diffuse reflectance surface (sintered Polytetrafluoroethylene (PTFE)) with a specific reflectance standard and a completely turned-off light source and blocked camera lens environment are used to take the white and dark reference images respectively (Williams et al., 2012).

It is paramount to note, that the choice of the value of the reflectance factor for the white panel could vary depending on the researcher's target and interest but once the geometry parameters are specifically defined, the HSI of the calibration sample will be certainly ready for caption at varying exposure time (Yoon et al., 2010; Gómez Manzanares et al., 2022). The most commonly used option is 100 %, but (Wang et al., 2014) applied a pixel–level calibration based on the reflectance values of 75 % spectral on the reflectance panel and a dark image to identify aflatoxin $B_1$ on maize kernel surfaces. While Khoshtaghaza et al. (2016) utilised a white reflectance tile with 20 % reflectance to provide an estimate of incident light on the tile at each wavelength and to calibrate hyperspectral images of rainbow trout samples. Invariably, with the assistance of white reference and dark images measured, absolute reference values of each pixel can be changed to relative reference values using the Eq. (1):

$$\mathscr{P}_{\mathscr{I}}(\lambda)_{(n,m)} = \mathscr{I}_{\mathscr{I}}(\lambda)_{(n,m)} - \mathscr{I}_{dark}(\lambda)_{(n,m)} \Big/ \mathscr{I}_{white}(\lambda)_{(n,m)} - \mathscr{I}_{dark}(\lambda)_{(n,m)} \times \mathscr{P}_{white}(\lambda)_{(n,m)} \tag{1}$$

Where $\mathscr{P}_{\mathscr{I}}(\lambda)_{(n,m)}$ signifies the calibrated reference values, $\mathscr{I}_{\mathscr{I}}(\lambda)$ referred to the hyperspectral image sample (measured raw value), $n, m$ pixel dimension, $\mathscr{I}_{white}(\lambda)$ referred to the HSI image of reference material (typically white) required to perform the spectral reflectance estimations, $\mathscr{P}_{white}$ is the spectral reflectance value of reference material, $\mathscr{I}_{dark}(\lambda)$ is the dark image required for the calibration of the sensor noise and background obtained by the image through the optical device. Hence, the white and black calibrations serves as data cleansing techniques required for adequate effectualisation of HSI outputs.

### 3.3. Throughput (illumination) correction model

High optical luminosity or throughput is an essential requirement in the optical design to prevent the effect of photon shot noise on the analytical signal. Gomez et al. (2020) revealed that the effect of illumination from a halogen source due to photochemical damage to artwork is of important caution in HSI system configuration. To this end, the choice of light-emitting diode (LED) may be considered as an alternative light source of illumination in HSI design. To resolve the non-homogeneous throughput problem in reflectance images, (Garcia et al., 2003) proposed a throughput–reflectance model. The model takes the image as a function of a product of the throughput and reflectance properties of a scene. Taking the camera characteristics into cognizance, which may cause gain and offset terms during the process of capturing images, the acquired image can be described by Eq. (2).

$$f(x, \quad y) = g(x, \quad y) \quad \times i(x, \quad y) \quad \times r(x, \quad y) \times o(x, \quad y) \tag{2}$$

Where $f(x,y)$ represents the image captured by the camera, $g(x,y)$ is the gain contribution of the camera, $i(x,y)$ is the throughput multiplicative factor, $r(x,y)$ is the ideal image under the absence of shading (reflectance function) and $o(x,y)$ is the camera offset. Since for a standard camera, the influence of offset is very small compared to $i(x,y)$, the value of $o(x,y)$ can be ignored. Thus, Eq. (2) can be expressed by:

$$f(x, \quad y) = \quad f_m(x, \quad y) \quad \times r(x, \quad y) \tag{3}$$

Where $f_m(x,y)$is the multiplicative factor that is produced by $g(x,y)$ and $i(x,y)$ and can be modeled as a function of smooth or morphological operation. To model non-homogeneous throughput, smoothing treatment or other morphological operations can be applied depending on the specific problems. The obtained image $f_m(x,y)$ can be treated as a reference standard to detect the effects of the illumination on each pixel. Then, the acquired image $f(x,y)$ can be modified by reference image in a point–by–point way, obtaining the ideal reflectance image $r(x,y)$ by:

$$r(x, \quad y) = f(x, \quad y)/f_m(x, \quad y) \quad \times \delta \tag{4}$$

Where $\delta$ is a normalisation constant that restores the overall image brightness. Above expression are of necessary for modelling to enhance illumination effectiveness during light source configuration for HSI smart system.

### 3.4. Operational principle of hyperspectral imaging (HSI) system

Consequent to the configuration of an HSI system, the samples brought for assessment be placed on the translation, rotating, or vertical stage while the system is switched on for image caption operation. The image acquired by the HSI device will then be stored in a microprocessor (i.e. Raspberry Pi) in the case of an embedded system for further image pre-processing otherwise be stored in a computer central processing unit (CPU). The image acquired by the HSI system consists of both the chemometrics properties (spectral information i.e. spectrum dimension at each point, wavelengths, texture, and 3D parameters) and spatial parameters (size, shape, colour, length and 2D signals). Collected images in the storage device will then subjected to the processing step as indicated in the image processing flow – chart (Fig. 6). The image pre-processing will then be conducted in three stages of analyses namely: low-level image processing, intermediate-level image processing, and high-level image processing. Each of the processing consists of unique multivariate algorithms to be deployed.

#### 3.4.1. Low-level image processing (LLIP)

LLIP is the manipulation of images acquired by a digital or HSI camera. Usually when digital cameras capture an image, the acquired sample can be transformed into digital form and processing such as image cropping, size reduction, brightness, redundancy and colour

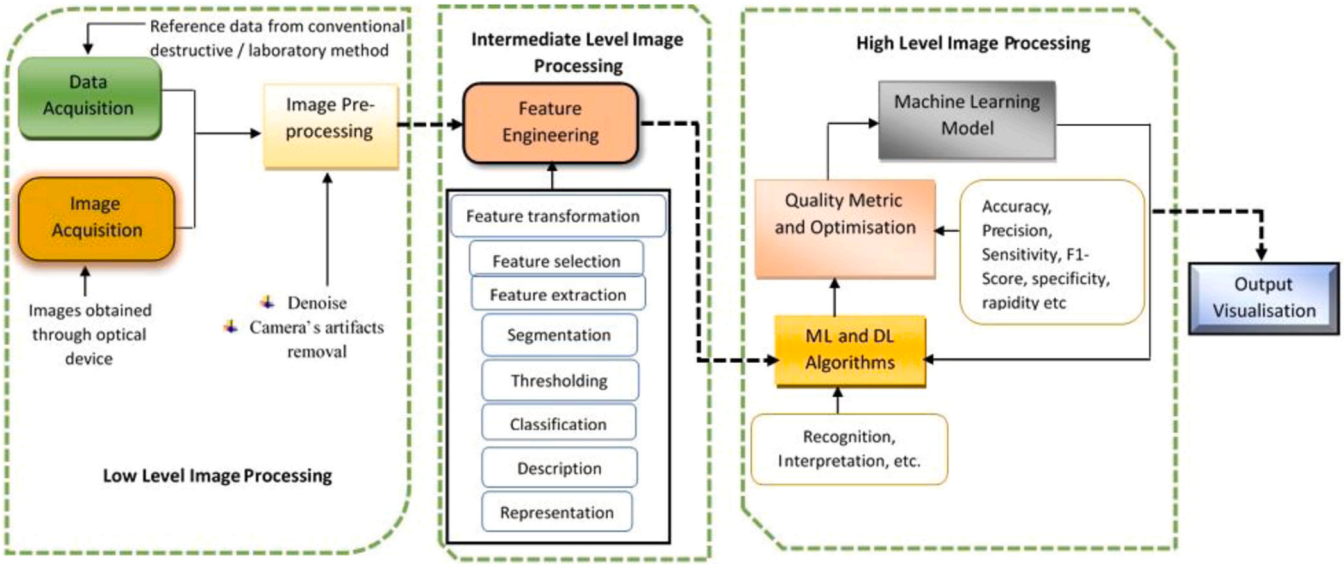



**Fig. 6.** Image processing stages for deep learning operations: ML – machine learning, DL – deep learning.

combination can be manipulated to produce the final image on the computer before image printing. In the case of HSI, once the simultaneous images are acquired, image pre-processing is conducted to remove the background noise, and eliminate the camera artifacts, insufficient light reflection, and other interferences to enhance the visibility of the ROI (Fig. 6). The images are loaded (from its save directory) to the current working directory for further conversions (grayscale mode, sample compression (resize), zero other mode (ZM), first order (FM), spiked points, dead pixels and ROI detection, etc) while numerous libraries (e.g. ipympl, imageio, matplotlib, numpy, scikit and skimage) have been employed to achieve these process using an open source software (PYTHON programing language) and MATLAB. The images portion of interest where both the chemometrics and morphometric properties of the sample are identified and are extracted. Mahendran et al. (2021) employed digital camera image acquisition and image pre-processing techniques to extract Acrylamide (AA), a carcinogenic substance in potato chips (PCs). The sample was placed at a distance of 50 cm from the camera while the digital camera focused the sample in the chamber of white fluorescent lamps fixed vertically to obtain the image samples. The authors applied the PYTHON module to conduct pre-processing activities (HSV, thresholding, and grayscale image conversion), image segmentation, ROI extraction and feature extraction methods in order to remove the redundant background colour of the sample. The algorithms for the AA quantification in terms of black colour ratio were developed using a PYTHON module with an Intel processor at the speed of 20.5 GHz, 8 Gb RAM, 64-bit Linux Operating system. The system accurately made the lowest prediction of AA content (0.26 μg/g) and observed the highest AA content to be 4.75 μg/g on the PCs in the experiment.

### 3.4.2. *Intermediate level image processing (ILIP)*

*3.4.2.1. Feature engineering.* Feature Engineering is the process of transforming raw data into features that are more informative, suitable and more relevant for machine learning (ML) or deep learning (DL) algorithm operations. It is a critical step in the ML and DL process as it can have a significant impact on the performance of the model. During the data pre-processing action, three common activities mostly conducted to improve the accuracy of HSI are: feature selection, feature extraction and feature transformation. The process essentially help to improve the image classifications and prediction accuracy. During this process, prior knowledge of the data being processed is required to ascertain the authenticity of the data generated by the algorithm. Therefore, mathematical methods which include fisher score, forward feature selection, and greedy expression are usually required to make the choice. Some suitable multivariate algorithms mostly adopted for conducting these actions are principal component analysis (PCA), independent component analysis (ICA), partial least square regression (PLSR), and many others. This process is referred to as intermediate-level image processing (ILIP) (Fig. 6).

ILIP is the process of subjecting extracted digital images to feature engineering techniques, image segmentation, image description and image representation algorithms for proper prediction. Image segmentation has been an important phase in image processing evolving the division of an image into constituent parts using a variety of methods (edge detection, region growing and thresholding) (Fig. 6). This process enables the separation of target objects from unwanted information which reduces the computational cost, and duration of image analysis and enhances image prediction accuracy. Momin, 2017 innovated MVS used for grading mango fruits. In the study, mango images were captured with colour camera, and image segmentation techniques were conducted which included thresholding, pattern recognition and deformable models were utilised to perform the segmentation exercise. Features projected area, sample perimeter, sphericity (roundness) and ferret diameter were identified accordingly to estimate the shape and geometrical properties of the fruits. Hence, the feature projected areas were used for sample filtering and pixel counts to grade the fruits into small, medium, and large categories. Hosseinpour. (2015) designed a novel rotation-invariant and scale-invariant image texture processing approach, to eliminate the effects of sample shrinkage on the detection of texture features during the monitoring of in-line shrimp drying. Image of texture features were fed into a Multilayer Perceptron-ANN (MLP-ANN), to predict the moisture level of shrimps.

The challenges of HSI spectral variabilities emanating from environmental factors (e.g atmospheric conditions such as humidity and local temperature) and background noise as a results of instrumental configuration (sensor and other components selection) obviously affect remote sensing (RS) output. However, to address the issues, application of spectral mixture model tagged augmented linear mixing model (ALMM) has played significant role in the field of geoscience. This was achieved by extensive application of synthetic and labeled datasets which demonstrated effectiveness and supremacy model as conducted by Hong et al. (2019) over previously existing state-of-the-art techniques. Numerous other pixels in hyperspectral information is also faced

with the effect of material mixtures due to lower spatial resolution generated from the camera configuration. Mixed pixels invariably degrades the performance of high-level data analysis. Hence, HSI spectral unmixing methods has surfaced to tackle the difficulties and has been deployed in variety of geoscience applications including mineral mapping, land cover changes detection (Bioucas-Dias et al., 2012; Rogge et al., 2006; Adams et al., 1995). Fu et al. (2016) proposed a spectral-library-based spectral unmixing approach, called the dictionary-adjusted nonconvex sparsity-encouraging regression (DANSER), to model the mismatch between the spectral library and the observed spectral signatures.

Thouvenin et al. (2016) indicated that spectral variability can be represented using a perturbed linear mixing model (PLMM), where the variability is explained by an additive perturbation term for each endmember. One drawback of the model is a lack of physical meaning. For instance, as a principal spectral variability, scaling factors should be coherent with endmember spectral signatures, while other variabilities are often incoherent with endmember spectral signatures. Intuitively, such attributed spectral variability cannot be represented by an additional term. In contrast, an interesting approach, called an extended linear mixing mode (ELMM), has been proposed in (Veganzones et al., 2014; Drumetz et al., 2015).

To curb the limitations of both ELMM (inability to explain scaling factors due to non-convexity) and PLMM (too generic to model various spectral variabilities), ALMM simultaneously modeled some scaling factors and other spectral variability based on distinctive properties by the endmember dictionary include Few-Shot Learning (FSL), Data Augmentation (DA), Semi-supervised learning (SSL), Generative Adversarial Network (GAN), Transfer Learning (TL), Transfer model (TM), light detection and ranging (LiDAR), Active learning (AL). Ensemble and Deep Ensemble Learning (Ullah et al., 2024). A synthetic data-driven dictionary learning method was explored in the framework to expand the scalability of the endmember dictionary. At last, ALMM was successfully endorsed to promotes the spectral variability by adequately harnessing an algorithm for learning the spectral dictionary. Hence, examining the experimental results on a synthetic dataset and two real datasets, this methods culminate into taking the spectral variability into considerable and superb algorithm compared to previous models, since the model separately handle spectral variability as scaling factors and other spectral variability based on its distinctive properties (Hong et al., 2019).

*3.4.2.2. Feature extraction.* Feature extraction is the transformation of real data into new destination feature space through projection pursuit (PP), minimum noise fraction (MNF), PCA, Orthogonal centroid algorithm (OCA), ICA and others for handling food and agricultural products. Morphological Profiles (MPs), including many texture features such as gray-Level Co-occurrence Matrix (GLCM), Gabor filters (GFs), Local Binary Patterns (LBP) and DNN-based algorithms are some popular techniques for the extraction of spatial information from HSI cubes. Geometrical properties can be extracted from morphological profiles. Some improvement in MPs involves Extended Morphological Profiles (EMPs) (Palmason et al., 2005), multiple-structure-element morphological profiles and invariant attribute profiles (Hong et al., 2020). During image classification and chemometrics analysis of samples, texture analysis which represents the spatial dispersion parameter of various image variations at respective wavelengths is often used to obtain physical characteristics of food samples including ripeness. For instance, Gamal Elmasry, (2007) employs GLCM to obtain the level of ripeness of strawberry classification which generated a higher level of classification accuracy of 89.61 %. This is a pointer to the effective utilisation of the whole process prescribed in Fig. 6.

### 3.4.3. High-level image processing (HLIP)

HLIP is the application of deep learning algorithms to extract vital information from a weighted sum to make classification decisions of target objects depending on the number of layers evolved and the complexity of the input variables to the choice of activation function. This provides the system with the ability to mimic the human brain and launch things with human configuration using different algorithms. Deep learning techniques consist of several frameworks depending on the target desires, some of the techniques are convolutional neural network (CNN), deep neural network (DNN), fully convolutional neural network (FCNN), recurrent neural network (RNN), binary feedforward network (BFN) or feedforward neural network, multilayer Perceptron (MLP), generative adversarial network (GAN). Other image processing and deep learning approaches specifically designed for hyperspectral imaging configurations include but not limited to hyperspectral anormaly detection (HAD), low-rank representation (LRR) model and autoencoders (AEs). An activation function is a mathematical operation in the deep learning process conducted to input a neuron (layers of neurons) in the network to introduce non-linearity to the network. Activation functions enable the network to learn complex patterns in the image data otherwise the entire network remains a linear combination of inputs, which results in limited modeling capabilities. The activation function allows the network to handle complex input analyses with ease of operation. Numerous activation functions exist in deep learning namely: Rectified linear unit (ReLU), Sigmoid, hyperbolic tangent (Tanh), Leaky ReLU, Parametric ReLU (PReLU), Softmax, Gaussian Error Linear Unit (GELU), Exponential linear unit (ELU), Scaled exponential linear unit (SELU), Softplus, Linearly scaled hyperbolic tangent (LiSHT), Swish and Average Biased ReLU (ABReLU) (Roy et al., 2023) (Fig. 7).

HLIP exercise can be demonstrated in Fig. 7. Some of the classification algorithms commonly used in machine learning operation are K-nearest neighbor (KNN), partial least square regression or discrimination analysis (PLSRA or PLS-DA), support vector machine (SVM and SVR), multiplicative scattered correlation (MSC), spectral angle mapper (SAM) and fuzzy logic (FL). Other multivariate algorithms that can be deployed include genetic algorithm, multiple linear regression (MLR) and standard normalised variate (SNV). Numerous algorithms have been implemented in food process applications for the development of machine vision system (MVS) but the choice of the algorithms depends on the researchers' interest and suitability. Ahn et al. (2019) deployed a deep learning (DNN) approach to particularly estimate the nutrient compositions (carbohydrate, protein and fat contents (CPF)) of selected food products using multiple deep neural networks. During the model architecture, the system utilised 3 multimodal DNN modalities to train the acquired images of the selected food samples (juice drink, meat, grains, snack and sausage). The modality used to build the system includes multimodal architecture between common networks (CN), nutrient-specific networks (NSNs) and an auto-encoder which consists of a CN and verification network (VN). The system was able to overcome overfitting difficulties in the training section by performing data augmentation that generates several similar data from each of the hyperspectral signals by adding random noises of small magnitude. Error quantification incurred during the operation was realised by means of utilising a symmetric mean absolute percentage error (SMAPE) which converts mean absolute error (MAE) into a percentage unit. As a result of the metrics, SMAPE gets closer to 0, the error becomes smaller. The system based on multimodal DNN had $R^2$ values of 0.9150, 0.7115, and 0.6047 for each of the food nutrients (CPF) respectively. The system demonstrated that DNN and multimodal DNN had similar performance and that the degradation of accuracy in both systems was mainly caused by outliers or erroneous predictions. To cap up the situation, the system successfully removed outliers, and hence, achieved $R^2$ values of 0.9543 (carbohydrates), 0.8527 (proteins), and 0.8481 (fats), and the average $R^2$ value was 0.885.

## 3.5. Hyperspectral data and image acquisition techniques

Data acquisition can be conducted through two different forms

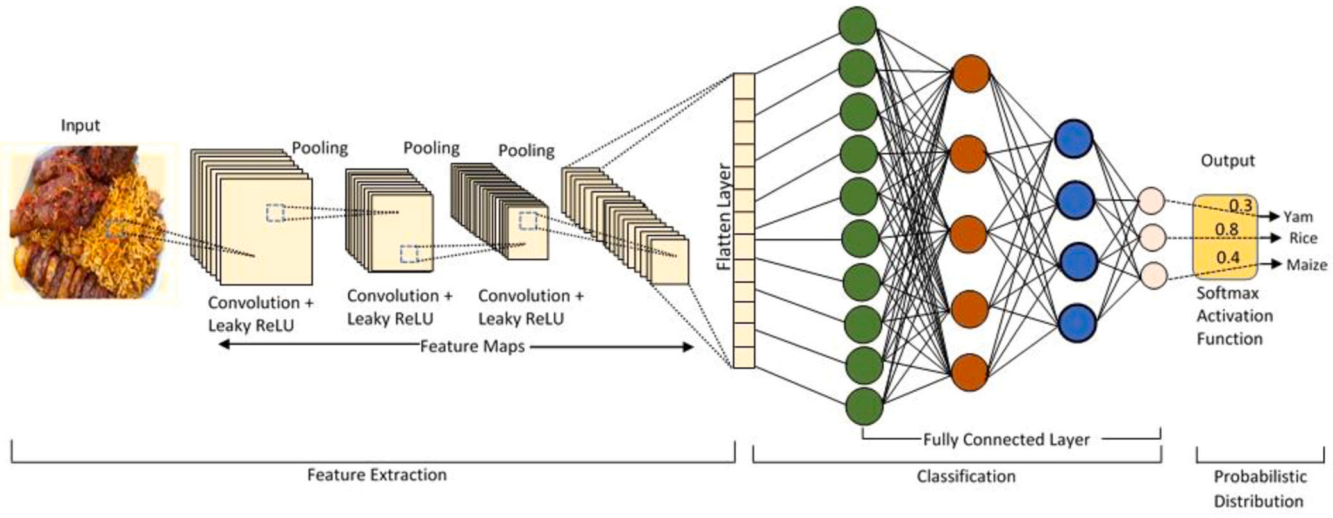



**Fig. 7.** Model architecture depicting deep learning operation: ReLU - rectified linear unit.

namely: through conventional laboratory analysis (destructive methods) where real sample nutritional facts will be quantitatively estimated to obtain label spectral parameters and secondly by using image acquisition techniques through hyperspectral imagery or any capable digital device (camera). HSI is known for its ability to simultaneously obtain three dimensions of images which comprise two spatial dimensions ($X_{axis}$ and $Y_{axis}$) and one other spectral dimension ($\tau_{axis}$) at each wavelength (Jia et al., 2020). When HSI is acquired from the camera device, the spectra of the different substances tend to display different curves which include background noise spectra, reflectance, and the real image of substances that are quite significantly different from one another. The images stored in the storage bank (CPU or Microprocessor) is further subjected to image processing. Based on the spectral signatures, the different materials can be identified and separated into different unique categories. The hyperspectral images acquired (hypercube) and fundamental operation for image collection of one spectrum for a single image are classified into four namely: pushbroom (line by line) imaging system, whiskbroom (point by point), liquid crystal tunable filter (area scanning) and full-waveband image at a time mostly denoted as snapshot (Fig. 8).

Among the highlighted four image acquisition principles, the snapshot is a non-scanning HSI that has no moving part and can record a full three-dimensional (two spatial and one spectral) hyperspectral data cube with each video frame. Although, pushbroom imaging system are the most commonly used acquisition mode for online applications in food industry and researches due to its versatility, rapidity and feature exploration potentials. For instance, Ahn et al. (2019) perform experiments using 140 real foods (juice, meat, grain, sausage and snacks). The HSI signals for all the food items were conducted using a short-wave IR (SW-IR) device (PANIMA, NIP Inc., Gyeonggi-do, Korea) with an IR lens (FA-megapixel, CBC Inc., North Carolina, USA). The IR lens can sense a wavelength range of 887∼1722 nm. The research captures SW-IR images from multiple regions of the food samples to produce its hyperspectral signals by averaging values of 5 different regions of the samples to ensure spectral stability. The system therefore was trained using the relationship between measured spectral signals and their associated label data (ground truth values). The outcome of the research demonstrated an accurately estimated CPF value (i.e. average $R^2$ value of 0.885 and SMAPE value of 0.1189), while effectively suppressing estimation errors. The outcome of the researchers' effort is a clear evidence that the point-wise hyperspectral signal is unstable since different food and agricultural produce consist of different absorbance, reflectance and electromagnetic energy at a specific wavelength due to chemical compositions and physical characteristics. It is also an indication that spectral stability may also vary since the ingredients of the food to be obtained differ from one region to another.

## 4. Machine vision system (MVS)

MVS is the process of training a computer-aided device (system) to mimic the human brain by using cameras to observe, obtain data (image information, time series and numerical data), and process the data in a much more accurate and rapid capacity to make adequate prediction (using some machine learning or deep learning algorithms). MVS utilises

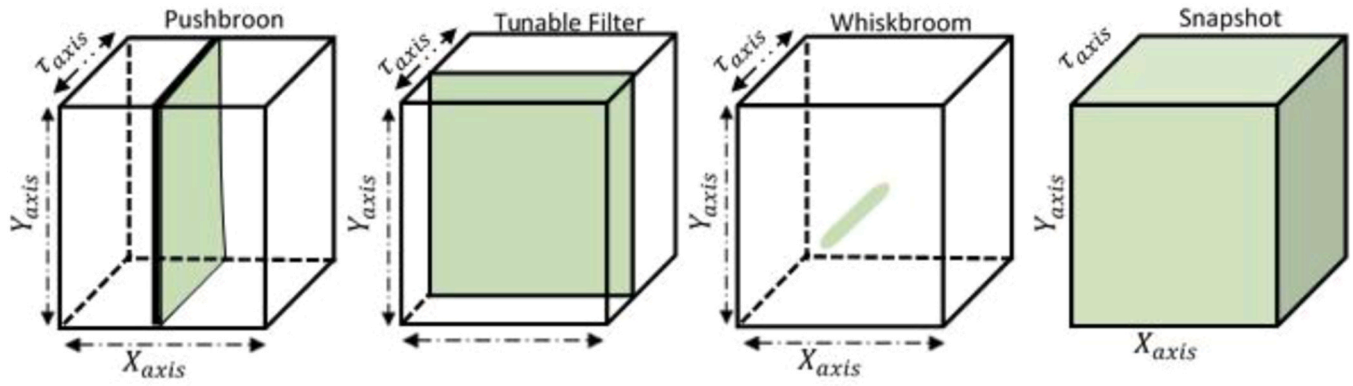



**Fig. 8.** Hyperspectral image data acquisition techniques.

deep learning (DL) and machine learning (ML) algorithms to execute the desired task. The system has been used to implement several research projects even on the factory floor. The main process of a typical MVS is described in Fig. 6. Sun et al. (2017) applied a $360^0$ rotating HSI system to detect fungi that cause peach decay at the peach surface. Likewise, Hu et al. (2016) and Fan et al. (2017) use HSI reflectance imaging techniques to automatically detect the bruise and damage of blueberries respectively. During the processes, the three levels of image processing (low-level image processing, intermediate-level image processing and high-level image processing) were embraced.

Despite improved computational cost being one of the major challenges of handling 3D data with respect to huge spectral parameters of various forms and reduction in inference efficiency of masked imaging modeling (MIM) which allows for the flexible use of various deep architectures as network backbones. Vision transformers (ViT) and Swin transformers has tremendously contributed to the potential of pretrained models for RS applications and rapid progress in image spectroscopy in Earth observations (EO) and can also be deployed in the processing of food and agricultural products (Liu et al., 2021). Even though, CNN-based HSIC approaches is able to accomplished significant advancement in terms of classification results, the need for HSI modification and integration to an embedded system is required where the framework are incorporated into a single unit. This is of utmost importance to ease spatial and spectral features for HSIC. Many of the methods discussed previously employ dimensionality reduction (DR) techniques to improve spectral spatial representation, but MIM, ViT, Swin transformers and other image processing skills integrated into a modified HSI technique will help to eliminate HSI spectral features due to system portability and potentials of the models embedded. It will also enhances the computation complexity and slows down the training procedure; consequently, parallel processing of networks utilising Graphical Processing Units (GPUs) and Field-Programmable Gate Arrays (FPGAs) will represents a pioneering approach to optimise computational performance while maintaining model efficiency. Concurrent harnessing the parallel computing abilities of GPUs and FPGAs can also enable the network operations to be distributed across multiple cores, allowing for the simultaneous execution of tasks and significantly reducing processing time (Ullah et al., 2024).

### *4.1. Machine learning (ML) and deep learning (DL) operations in a smart system*

Both ML and DL are all subset of artificial intelligence (AI) tools used for the configuration of MVS relationships as displayed in Fig. 9. Consistent application and development of ML approaches have significantly reduced the challenges and pressure of data analysis for ease of model prediction. To this end, DL techniques have been developed to handle complex data (in terms of image) manipulation due to their complex architecture possession and improved data analysis capacity. Substantial efforts has been deployed specifically toward enhancing computer vision system representation performance for hyperspectral anomaly detection (HAD). HAD relies on statistical manipulation via physical model-based methods and deep learning-based techniques. Among the commonly adopted methods, low-rank representation (LRR) model are widely used for its formidable competences on background and target interest, however, LRR-Net influences the alternating direction method of multipliers (ADMM) optimizer to resolve the LRR model imperfection by incorporating prior knowledge solution (i.e reference data) into deep network to control optimisation of parameters. HAD also displayed very powerful resolution for the optimisation problem with sufficient accuracy, interpretable and fast convergence rate in HSI deep network architecture (Li et al., 2023a,b). Other interesting deep learning techniques specifically designed for HSI enhancement include but not limited to Deep Belief Network (DBN), autoencoders (AEs), generative adversarial networks (GANs), recurrent neural network (RNN) and long short-term memory (LSTM) (Wu et al., 2023; Yang et al., 2019; Wang et al., 2022; Jiang et al., 2020; Lyu and Lu, 2016; Kieu et al., 2018). While other categories of researchers generally combine the AE and GAN to yield an adversarial generative network (AAE), all in an efforts to ensure technological simplicity for handling HSI for various substances useful for mankind (Lu et al., 2020; Jiang et al., 2020; Xie et al., 2021, Roy et al., 2021).




**Fig. 9.** Embedded Artificial intelligence system principle. AI – artificial intelligence, CNN – convolutional nueral network, RNN – recurrent neural network, LSTM – long short-term memory, SVM – support vector machine, RF – random forest, LR – logistic regression, K-nearest neighbor.

Furthermore, a pseudolabel-based unreliable sample deep learning (PUSDL) technique for handling semi-supervised HSI data classification has emerged indicating promising excellence performance in the model architecture. The model uniquely select low-confidence unlabeled samples containing abandon large number of rich land-covers information in contrast to the pseudolabel-based techniques which utilises high-confidence unlabel data (Yao et al., 2023). PUSDL was adopted to avoid overfitting the spatial distribution of the samples for building position – free transformer (PTF) as the backbone classification networks. Hence, unreliable unlabeled samples are treated as negative samples for the corresponding categories to improve the discrimination of PFT in a contrastive learning model. The study successfully justify the use of numerous reliable unlabeled and unreliable unlabeled samples for training HSI model and guaranteed it efficiency. Other team of researchers deployed a single PC techniques to train a spatially-based 2D-CNN model by incorporating preprocessed PCA whitening limited to three principal components (PCs) into randomized patches network for HSI object classifications (Haut et al., 2019). Even though, small amount of labeled data for training instances with relatively similar spectral features can cause DL models to be sensitive to the overfitting problem. But to tackle such experience, Wang et al. (2022) offered an innovative approach called Probabilistic Neighborhood Pooling-based Attention Network (PNPAN) specifically designed for HSIC (hyperspectral image classifications) some introduced a Deep Feature Fusion Network (DFFN), heterogeneous kernel convolution (HetConv3D)

specifically enhancing HSIC processing (Song et al., 2018; Roy et al., 2021).

#### 4.1.1. Traditional machine learning techniques in food processing and postharvest technology

Traditional machine learning (TML) is the application of a computer to extract useful statistical features on sample datasets (numerical or label data) to make valuable model predictions by providing a framework for solving the learning problem when limited datasets are involved. TML can be classified into supervised machine learning (contains label, known or past data), unsupervised machine learning (contains unstructured data or images) and reinforcement (contain some label data while generating some for optimisation). The main analyses conducted in TML are regression analyses, pattern classification probability density estimation, and many others analyses. Related algorithms that are commonly used here are support vector machine (SVM), random forest, logistic regression, K-nearest neighbors (although known as lazy classifier algorithms), K-means cluster, Bayesian network, Fuzzy C-means and Decision Tree (Fig. 9). TML has been deployed for many states-of-the-art project implementation. Commonly used feature extraction techniques mostly adopted in ML operations are principal component analysis (PCA), wavelet transform (WT), independent component correlation algorithm (ICA), scale-invariant feature transform, speedup robust features, histogram of oriented gradient and many others (Zhou et al., 2019). Aghbashlo et al. (2014) in a study applying the implementation of intelligent systems to monitor and control food drying speed to improve the quality of dried food substances while maintaining the nutrient content. The study inspected features (texture, colour, size and shape) of the food using GLCM, Fourier transforms (FT), Wavelet transforms (WT), thresholding and masking the food images (input data). Subsequently, PCA and FCM were applied to control the moisture contents in the process to attain optimal dryness.

Food is any edible substance consumed for its nutritional benefits and provision of energy for the body system. Postharvest technology on the other hand is the techniques for the safe handling of food produce from the farm gate to the consumption level without damaging the nutrient contents of the food sample. Therefore, food processing is the conversion of raw food materials into new edible substances (form) while maintaining its nutritional quality. Food processing is classified into primary (rough) processing and secondary (deep) processing depending on the unit operations involved. Primary processing is the initial unit operation conducted on agricultural produce after harvesting, to maintain the original nutritional compositions of the produce from being damaged, or lost and to meet storage, transportation and future processing requirements (Zhu et al., 2021). Some of the techniques involved in primary processing include milling, pasteurisation, shelling, drying, salting, freezing, slaughtering, frying and tempering. Secondary (deep) or postharvest processing is a more elaborate unit operation technique conducted to transform food produce to a final product while improving the feature characteristics of the products following the primary operation. For instance, cereal grains such as rice, and wheat can be processed into multigrain fermented fried batter and snacks (known as "*masa*" and "*sinasir*" respectively in Northern Nigeria), grains can also be converted into bread, biscuits, snacks, noodles, and couscous. Irish potatoes, yam and cassava among the root tuber crops can be processed into chips, pounded yam and cassava flakes ("*Gari*") respectively, while fruits can be converted into juice drinks and various wines. Deep processing is an essential aspect of food management and postharvest technology which helps to enhance food security, prolong food shelf-stability, increase the economic values of agricultural products and contribute to human diet.

Traditional (conventional) food processing methods are time-consuming, labour-intensive, and are prone to error while a constant increase in human population across the globe and diversification of consumer demands has created significant pressure on the food industry and research institute. This is due to operational costs, material wastage along the production line and lack of technical know–how. Therefore, these processing steps cannot constantly yield high quality, low price of commodity and rapid response to global demand (Capitanio et al., 2010). Despite, the traditional techniques are still demonstrating an important role in both primary and deep processing unit operations in the food industry. Researchers are constantly innovating and establishing informed decision-making and emerging food processing techniques to reduce the pressure, reduce the cost of operation, enhance food quality and improve processing efficiency (Van Der Goot et al., 2016). Currently, state-of-the-art equipment can be employed in all the unit operation stages in the food production and supply chain (from farms – to factories – consumption level) (Manzini & Accorsi, 2013). Technological advancement has led to MVS development utilised for many unit operations including canning, freezing and drying processes to prolong food preservation and storage duration, adequate packaging process and detection of foreign objects in food substances (Langelaan et al., 2013). Along the processing chain, some auxiliary protocols such as high-quality food identification and sub-standard foods are observed. The flow of operation includes food safety and quality evaluation, food preservation, food processing monitoring and packaging, and foreign object detection. In this way, it is possible to effectively prevent the waste of resources and improve productivity. Appearance, texture, and components of food are the primary references for determining the quality and safety of food, monitoring food processing procedures, and detecting foreign objects (Cardello et al., 2007; Ma & McSweeney, 2008). Similarly, Zareiforoush et al. (2016) designed an intelligent system using MSV and Fuzzy logic techniques to control rice whitening process. The system controlled the process by analysing the degree of milling and the percentage of broken kernels. The result of the fuzzy logic algorithm (that is the whitening pressure level) can be deployed to realise the process monitoring. The authors ascertained that the system's efficiency was 31.3 % higher than manual labour with an accuracy of 89.2 %. This indicates that an informed decision is faster, less intensive and can reduce drudgery.

#### 4.1.2. Deep learning techniques (DLT) and algorithms

Deep learning (deep neural network (DNN)) is the learning technique used for the manipulation of images to build a deep architecture and for generating models by iterating some activation functions in multiple layers. The ability of deep learning to resolve complex analyses makes the learning more efficiently robust and superior to the TML methods although the TML is more efficient in terms of interpretation ability, lance and accuracy. Deep learning algorithms (DLA) are powerful mathematical algorithms that are based on representational learning of data (images). The tools can be used to observe edges, region of interest and intensity of images pixel by pixel. Unlike the ANN (used for numerical data) framework, these methods have been applied to solve many computer vision difficulties such as speech and face recognition, pattern recognition, autonomous vehicles, natural language processing and bioinformatics which have yielded excellent outputs. Other models used for deep learning operations are long short-term memory (LSTM), stacked autoencoders (SAEs) whose structure is related in convolutional ANNs and single-short multi-box detection (SSD) (Zhou et al., 2019). Several handcrafted-based feature extraction (FE) and classification methods in HSI have utilised various texture descriptors, some include local binary patterns (LBP), histogram of oriented gradients (HOG) (Dalal et al., 2005), global image scale-invariant transform or global invariant scalable transform (GIST), pyramid HOGs, scale-invariant feature transform (SIFT) (Oliva & Torralba, 2001; Lowe, 1999).

SIFT is a broadly applied dynamic feature descriptor utilised for computer vision applications (Azhar et al., 2015). The limitation of SIFT descriptor was insensitivity to alteration in image scale, illumination, background noise, rotation and computational intensive. SIFT's raises computational complexity concern which was later overcome by other approaches. Then scales and orientations (gradient features) of unique

sub-regions of an image are characterized by GIST, which is the global description of crucial factors of an image. GIST generates a spatial framework based on countless statistical characteristics, such as openness, roughness and ruggedness (Huang et al., 2016). Texture descriptors such as local binary patterns (LBPs) are applied for remote sensing image analysis (Nhat & Hoang, 2019). By selecting pixels from the square neighborhood, LBPs are employed to define the texture around each pixel. The grayscale scores of all neighboring pixels are threshold concerning the central pixel. The color histograms being an important action in image processing has adopted GIST descriptors, and various texture representations to extract global image features that characterise statistical properties such as textural patterns, color distributions and overall spatial structure. In contrast to the local feature descriptors identical to HOG and SIFT capture geometrical information from local neighborhoods. These local descriptors are often aggregated into bag-of-visual-words (BoVW) models and HOG feature-based models (Ullah et al., 2024).

To improve the constraints and limitations of the handcrafted feature extraction, a deep feature learning technique was presented by Hinton and Salakhutdinov (Hinton & Salakhutdinov, 2006). DL-based methodologies are applied to extract features from data in a hierarchical fashion to build a model with progressively increasing semantic layers. This is accomplished once the model acquires a suitable representation of the data. Such models have demonstrated promising effects for feature representation in HSIC (Hu et al., 2015).

#### 4.1.3. Artificial neural network (ANN)

ANN is a relatively simple computational model structure inspired by the functionality of biological neural networks in the human brain. The simple intuition uses a mathematical function to obtain the weighted sum of inputs and outputs its results. This framework usually contains three layers namely the input layer, the hidden layers and the output layer. The hidden layers mostly consist of interconnected neurons (nodes). The input layer receives a signal (input or initial data) of the external world to be processed. The hidden layer situated between the input and output layer conducts computations on the input data and converts it to other features that can be quantified with the help of a weighted sum to generate an output signal through the output layer (Zurada, 1992). Along the learning process, ANN performs remarkable tasks by adjusting the weighted sum between the neurons through a process called back-propagation. Back-propagation evolves by feeding input data forward through the network while comparing the predicted output to the actual output. In the process, it estimates the error and then propagates the error backward through the network to adjust the weights and minimise the error for optimisation. This process helps to evaluate the gradient of the loss function for weight in the network (Hecht-Nielsen, 1992).

#### 4.1.4. Convolutional neural network (CNN)

CNN is a specialised type of ANN that consist of a series of convolutional layers, nonlinear layers, pooling (down-sampling) layers and fully connected layers concerning it activation function specifically designed for processing structured grid-like data (images) (Krizhevsky et al., 2012). In CNNs, the convolutional layers apply a set of learnable filters (kernels) to the input data while each of the kernels detects specific features in the input such as edges, textures and shapes. CNNs are predominantly effective in tasks related to image recognition and classification, object detection, and image segmentation. The output image can be interpreted as several matrices and the output is the determination of what object the picture is most likely to depict. In CNN several activation functions can be applied to the output of convolutional layers to introduce non-linearity into the network and enable it to learn complex patterns. The result is used as the output of the layer. Pooling layers usually down-sample feature maps produces by convolutional layer by aggregating information from neighboring pixels. Pooling helps reduce the spatial dimensions of the data and makes the representations more invariant to small translations and distortions. Fully connected layers are typically placed at the end of the CNN architecture and perform high-level reasoning on the extracted features (Fig. 8). Each neuron in a fully connected layer is connected to all neurons in the previous layer, effectively flattening the feature maps into a vector representation. The filter will slide across the entire image and repeat the same dot product operation. Fully connected layers are often used for classification tasks. The process of convolution is expressed as given in Eq. 5.

$$s(i,\quad j) = (X^{*}W)(i,j) + b = \quad \sum_{k=1}^{n_in} (X_k{}^{*}W_k(i,j) + b) \tag{5}$$

Where $n_in$ is the number of input matrices or last dimension of the tensor, $X_k$ represent the $k^{th}$input matrix, $W_k$ is the $k^{th}$ sub-convolution kernel matrix of the convolution kernel and $s(i,j)$ is the value of the corresponding element of the output matrix related to the convolution kernel $W$. For the output of non-linear layers, the ReLU activation function $f(a) = \max(0, z)$ is generally used. The ReLU function returns a value of 0 for each negative value in the input image but returns the same value for each positive value in the input image. The sigmoid activation function maps any real-valued number to the range of (0, 1) which enables the useful binary classification prediction. The formula for the sigmoid function is expressed as given in Eq. 6.

$$\delta(x) = 1/1 \quad + \quad \varepsilon^{-x} \tag{6}$$

Where: $\delta\,(x)$ is the output of the activation function, $\varepsilon$ is the base of the natural logarithm and $x$ is the input to the network (function). However, the softmax activation function is frequently utilised in neural networks, especially in the case of multi-class classification tasks. It usually takes a vector of arbitrary real-valued scores (logits) and normalises them into a probability distribution over multiple classes. This makes it suitable for problems where the output requires to represent the probabilities of each class as the formula for the softmax applied to a vector $Y = (Y_1,\ Y_2,\ \ldots,\ Y_n)$ is given in Eq. 7.

$$Softmax(Y)_i = \varrho^{Y_i} \Big/ \sum_{j-1}^{n} \varrho^{Y_j} \tag{7}$$

Where $Softmax(Y)_i$ is the i-th element of the output vector after applying the softmax function, $\varrho =$ base of the natural logarithm $Y_i$ is the i-th element of the input$Y$, and $\sum_{j=1}^{n} \varrho^{Y_j}$ is the sum of the exponentials of all elements in the input vector$Y$. Softmax activation function fundamentally receives the exponential of each input score, which guarantees that the resultant output probabilities are non-negative, and then normalises them by dividing by the sum of all exponentials. This will therefore ensure that the output probabilities sum up to 1. Overall softmax is a fundamental component in the output layer of neural networks used for multi-class classification tasks.

There are numerous popular CNN architectures used for image processing and manipulations which include AlexNet reported by Krizhevsky et al. (2012), a network using a repetitive unit called visual geometry group network (VGG), GoogLeNet used by Szegedy et al. (2015), this includes parallel data channels, and residual neural network (ResNet) constructed by residual blocks (He et al., 2016). Additionally, these highlighted architectures can be obtained from the model reserve-park with pre-trained weights. This means the models have already been trained by some image datasets like ImageNet so the pre-training models must have learned the ability to extract image features (such as colour, texture information, high-level abstract representations and so on) (Deng et al., 2009). The itemised retraining method is called “fine-tuning,” which is proved to be a successful method to shorten the training time and to gain more accurate result. Convolutional networks have also been widely used in food or non-food classification, food category discrimination, and ingredients identification. These models has been deployed in the development of an APPs for

the identification of food, its qualities and quantification of food calories. For instance, Myers et al. (2015) established an Im2calories mobile vision automation APP for the detection of food diary calories and depth volume of food from images. The system was based on a 5-step training processes. GoogLeNet deep learning model was first deployed to fine-tune the CNN architecture and hyper-parameter. Secondly, the fine-tuned architecture was later employed to identify and recognise the food sample through the image caption. The third stage was conducted for semantic image segmentation to enable the system to obtain the food location. The fourth action was the prediction of the physical size of the segmented food. Lastly, the total calorie constituent and set of food samples present in the restaurant database were estimated using Eq. 8 based on the detected food samples their calorific density and volume accordingly. The architecture used allowed the system to parse the image of the food into a list called $\kappa$-items and Eq. 9 was used to estimate the set of food present by picking the suitable threshold $\varphi$ (probability of food sample seen in the image higher than a certain threshold $\varphi = 0.5$, which implies the number of items satisfying certain conditions, and $k$ indicates the index). An alternative means of calculating the calories value by automatic thresholding was also conducted to avoid bias in the study using (10).

$$\mathscr{C} = \sum_{k \in \mathscr{S}} \mathscr{C}_k \tag{8}$$

$$\mathscr{S} = \langle \mathscr{K} : \rho(\gamma_k = 1|x) > \varphi \quad \rangle \tag{9}$$

$$\widehat{\mathscr{C}} \quad = \quad \sum_{k=1}^{\mathscr{K}} \rho(\gamma_k = 1|x)\, \mathscr{C}_k \tag{10}$$

Where $\rho(\gamma_m = 1|x)$ estimates the probability that $\kappa^{th}$ the item seen in the image $x$ and $\mathscr{C}_k$ denotes the calorie constituents of menu item $k$. The enumerated sequence achieved a decent effect of mean absolute error (MAE) and mean error (ME) of $163.43 \pm 16.32$ and $31.90 \pm 28.10$ respectively.

#### 4.1.5. Fully Convolutional Network

Unlike the traditional CNN, a Fully convolutional network (FCN) is a unique kind of neural network architecture predominantly used in the field of computer vision systems to handle tasks such as semantic segmentation where the output is a labelled picture obtained by classifying each pixel in an image into different classes or categories (Long et al., 2015). Popular architectures based on FCNs include SegNet, U-Net and DeepLab. FCNs extend the capacities of CNNs to handle tasks like object detection and image generation by replacing the fully connected layers with convolutional layers. This allows FCNs to accept input images of arbitrary sizes and produce output images with pixel-wise predictions, maintaining spatial information throughout the network. During the operation, images or input up-sampling layers (transposed convolutions or deconvolutions layers) increase the spatial resolution of the feature maps, allowing the network to produce output at the same resolution as the input image.

### 4.2. Deep learning applications in food safety and quality evaluation

Eating habits and dietary intake often have significant effects on human health. Especially for hypertensive patients, diabetics, allergic folks and people placed under some dietetic watch. They required strict monitoring and control of their dietary behaviour to maintain sound fitness. Food identification, recognition and classification are essential tasks to help human beings record their daily dietary requirements. Food images are one of the most important pieces of information which reflect both the physical and chemometric characteristics of food items. Additionally, image-sensing devices are relatively effective detecting and safe acquisition tools for food appearance analysis. For climacteric, non-climacteric and processed food, there are large variations in food morphological properties such as shape, colour, texture, volume, and other physical compositions that make food and agricultural product recognition a challenging task. It is also a serious task to protect food samples during unit operation from cultivation through production to consumption due to several factors involved along the production chain. Damage to food products due to pests and disease attacks during and after production is another important concern demanding adequate attention to reduce losses of food and agricultural products during each unit's operations. However, several backgrounds and natural arrangements of food substances also introduce variations in the food recognition and classification process. Currently, due to the common use of CNN, image analysis has been the most frequently used pattern in food recognition and classification. Zareiforoush et al. (2015) conducted a study on computer vision and metaheuristic classification approaches to discriminate milled rice grains. The techniques adopted entail the extraction of certain features such as shape, texture, size and colour to build a primary feature vector and mine the most important features based on "greedy hill-climbing" and backtracking algorithm. The final feature vector was then utilised to train the SVM, ANN, BN (Bayesian Network), and decision tree (DT) independently. The outcome revealed that deep learning architecture ANN had the best classification accuracy of 98.72 % over the other multivariate algorithms. Wan et al. (2018) demonstrated the use of MVS to acquire tomato images to detect three maturity levels (red, green and orange) and the ROIs were segmented from the whole image sample. A sigmoid activation logistic function was applied while a Back-propagation Neural Network (BPNN) was used to classify the maturity level of both tomato varieties (Roma tomato and pear tomato).

Shafiee et al. (2014) used computer vision system (CVS) as a non-destructive tool to capture and perform the characterisation of honey based on colour transformation of the images and its correlated chemical attributes, while they used ANN to predict key quality indices of honey (ash content (AC), antioxidant activity (AA), and (TPC) total phenolic content). The system was able to convert RGB values to CIE $L^*a^*b^*$ colourimetric parameters with low generalisation error of 1.01 $\pm$ 0.99. He et al. (2023) in a study employed a rapid and non-destructive HSI to evaluate the chemical compositions (moisture, protein and ash contents) of chicken flesh by pre-processing the raw spectral using PLS built with Savitzky-Golay convolution smoothing (SGCS) spectral (SGCS-PLS model), MSC, BC, SNV and normalisation (NOR) methods to extract spectral parameters related to the chemical constituents. Four different techniques including the regression coefficient (RC), stepwise regression (SR), successive projections algorithms (SPA) and competitive adaptive reweighted sampling (CARS) were later selected to obtain optimum wavelength for sampling the original PLS models. During the study, 13 optimal wavelength were selected by correlation coefficient of calibration ($R_C$), the raw spectra (for protein) and RC-RAW-PLS model was constructed with correlation coefficient of prediction ($R_P$) of 0.923 and root mean square error of prediction (RMSEP) of 0.588 %. The overall outcome of the study indicates that HSI data combine with PLS algorithm can potentially be used for predicting the chemical compositions of chicken flesh.

Zhu and Spachos (2020) also developed a two-layer image processing system based on machine learning for banana grading and spot detection on banana peels, with an overall accuracy of 96.4 %. A support Vector Machine (SVM) classifier was first deployed on the banana samples to extract feature vector including the colour and texture. The second layer ($YOLO_V3$) meaning you only look once the version 3 model was later used to further locate the area of the defect on the peel and determine if the inputs belong to a mid-ripened or well-ripened class. The performance of the first layer achieved an accuracy of 98.5 % and the accuracy of the second layer is 85.7 %. This testifies that optical device used for machine vision system are capable of easily obtaining defects and other evaluations in food and agricultural materials based on the algorithm deployed for training the system and also the system sensitivity.

Rapid expansion of artificial intelligence potentials currently addressing global demand but require an updated techniques in the field of visual demonstration learning in a self-supervised manner. This is due to insufficient reference data in various specialisation coupled with the challenges and difficulties of handling large or complex spectral analysis. Researchers, recently discovered some technicalities of seamless data mining to enhance spectral remote sensing (RS) efficiency. Trending pre-training agents models adequate for spectral RS application in various field of application is divided into contrastive learning for pre-training (distinguishes between related and unrelated models) and generative learning pre-training (generate novel information or recover comprehensive data from partial observations) operation (Hong et al., 2024). Interestingly, continuous technological exploration is evident in the province of computer vision system (CVS) and natural language processing (NLP) leading to the discovery of spectral generative pre-trained transformers (SpectralGPT) and ChatGPT respectively. SpectralGPT been first of its kind "foundation model", recently demonstrated superb flexible ability to explicitly and efficiently resolve unique and sequential characteristics of millions of spectral parameters in remote sensing (RS) scene using masked autoencoder (MAE) structure with a modest and yet effective 3D GPT network. Recently, contrastive learning which fully focused on exploration of unlabeled samples for data pre-training has been extensively adopted in HSI classification specifically based on constructing high-quality positive and negative data samples pairs for learning discriminative features (Li et al., 2022; Huang et al., 2022; Chen et al., 2020). Researchers' in a study, firstly flipped hyperspectral image patches both horizontally and vertically, and then randomly detached the spectral parameters of non-central pixels in HSI patches to build different sample pairs (Hu et al., 2021). These contrastive learning-based means elaborately designed different data augmentation strategies to construct positive and negative sample.

SpectralGPT has further unravel and influence the potential of spectral remotes sensing data thereby providing in-depth ability to resolve several challenges and obstacles in the domain of spectral RS data analysis. Combination of fMoW-S2 +BigEarthNet used for pre-trained dataset has shown an exciting and superb outcome for the spectral analysis on Earth and environmental observation signifying considerable potential in progressing spectral RS big data applications within the field of geoscience across a number of downstream tasks including: semantic segmentation, single/multi-label scene classification, and change detection (Hong et al., 2024). Certainly, these advance deep network will enhance rich insights into the composition of food objects observed under hyperspectral imaging system building a transformative technology with vast potential to addressing global challenges in food security and reshape various industries.

### *4.3. Application of hyperspectral in food safety and quality evaluation*

Proper handling along the production chain and hygienic consumption of food products has been a major concern of consumers worldwide. This is vital to humans as it helps to guarantee consumer safety, provide information on the quantity of nutrient constituents and some level of confidence to people on the safety of food consumed at any particular point in time to avoid food contamination, deterioration and food poisoning. The traditional technique of these various assessments are destructive as they require some laboratory reagents and calculation which are strenuous and evolved sample wastage, thereby contributing to losses of samples along the chain. Hence, researchers, the food industry and consumers required a rapid and non-destructive means to evaluate the safety and quality of food products. In the food industry, morphometric features of meat and poultry, fruits and vegetables are related to the price of the commodities. Therefore, any physical changes in the food substance can reflect in the intrinsic chemical state which may result in product deterioration, unpleasant appearance, unpalatable and total sensory rejection (Zhu et al., 2021). Hyperspectral imaging system are however, able to detect undesirable changes which cannot be easily identified in foods by human eyes and that do not meet the consumption specifications, thus eliminating foods that are not suitable for international market standards. For example, variation features such as colour disparity, moisture content and texture are frequently used to assess the quality, safety and overall acceptability of poultry, meat, fruits and vegetable products for consumption. Numerous algorithms which is not limited to Gray Level Co-occurrence Metrix (GLCM), Local Binary Pattern and Gabor filter can be employed for such features. ElMasry et al. (2007) aspplied GLCM textural analysis to identify the ripeness stage of strawberries and partial least square (PLS) model to quantitatively predict total soluble solid (TSS), moisture content (MC) and level of acidity (pH) in strawberries. Dixit et al. (2021) used partial least squares regression (PLSR) and box plots to visually evaluate and predict the intramuscular fat content (IMF), pH and MIRINZ shear force values of beef. The observed result in the experiment has demonstrated that HSI Snapshot cameras and online Vis-NIR spectrophotometers are capable to offer an advantage in terms of rapid data acquisition. The study also provides supporting evidence for different types of practical implementation for meat assessment using destructive (contact) and non-destructive (non-contact) AI tools for the food samples evaluations. Hence, global acceptability of deploring new technological advancement in the assessement of food quality measurements are highly gaining recognition and is encouraged for adoption in the developing countries across the world to enhance food security, rapidity, precision and reduce drudgery along the production chain.

## 5. Summary

This manuscript reveals necessary components required to easily develop hyperspectral imaging (HSI) camera anywhere across the globe. The HSI being an artificial intelligence system is a trending technology used for handling various food and agricultural materials non-destructively. High-cost implication of HSI configurations mostly hindered researchers, institutes and professionals' accessibility to the promising optical device, hence the need for consistent efforts for seamless replication of the smart device availability is required. The manuscript also suggested possible modifications on the HSI system which includes integration of computer processing unit (CPU), high dimension multimedia interface (HDMI) monitor for output display and image acquisition device required for improvement on the existing system; thus, making it an embedded system to ensure ease of deployment of all necessary multivariate algorithms, portability and computational flexibility. The article summarised deep learning techniques and most currently adopted algorithms specifically utilised for HSI and other machine vision system, ranging from food and agricultural realm, environmental sanitation, aeronautics, surveying, transportation, healthcare, defense to communication.

### *5.1. Current development status and an outlook on future development trends*

Rapid development of artificial intelligence (AI) potentials currently addressing global demand and HSI is not an exception of both machine vision system and remote sensing technology. HSI as AI system is currently playing significant role as a non-destructive device thereby enhancing postharvest technology in food and agricultural industries. Many researchers have deployed numerous deep learning algorithms feature extraction and images processing techniques specifically for hyperspectral imaging configuration yielding promising output (Othman et al., 2023). Researchers have adopted advance segmentation techniques in deep learning classifiers, feature extraction, image processing and object recognition for handling various foods including fruits, vegetables and medical image classification which has helped to reduce processing time and improve detection accuracy (Jiang et al., 2023; Wang et al., 2018). Mostly, investigation has revealed that there was no single model capable of resolving all challenges encountered in

RS and MVS. Thus, comparative strategy adopting multiple algorithms are widely used for model optimisation in the classification architecture where the best are chosen through the algorithm performance of the model visual output (Gholian-Jouybari et al., 2023). However, a separate compartment in the existing HSI which necessitates modification of the system where integration of the graphic users interface (GUI), image acquisition device and central processing unit (CPU) is to be developed as a single embedded system. This will ensure ease prediction of food and agricultural material quality. Therefore multiple deep learning algorithms with potentials for significant improvements in the efficiency of image processing and mining features for direct deficiency detection in foods and other materials require more investigation on the system to ensure quality of food safety on the production floor (Zhou et al., 2019). While the application of DL model has demonstrated commendable potentials, little research has been conducted in the area of cooked food products and adulteration detection. Thus, more research is required to bring this goal to actualisation. The distinction between evaluation of food and agricultural material detection using non-destructive measurement of smart devices and destructive methods was elucidated using deep learning methods. This area of investigation requires adequate exploration using sophisticated AI models to increase focus in the future, using learning algorithms.

## 6. Conclusion

This survey conspicuously encapsulates current trend techniques deployed for the development of numerous hyperspectral image systems as a non-destructive artificial intelligence (AI) mechanism and its application in postharvest handling of several food and agricultural products. Invariably, the HSI system is currently gaining acceptance for non-destructive analyses and inspection of food products within the research arena and industries due to the total elimination of chemical reagents in the food sample's quality assessment (detection). The wider acceptability of HSI usage is also due to its rapidity, reliability and feature exploration potential. The informed decision by AI tools for food quality detection is an indication that the technology has shifted from biomedical, astronomical and other engineering applications and now been integrated into an effective tool for the exploration of quality assurance of various food and agricultural produce. Despite the existence of little challenges and the implementation of machine vision systems in food processing is a universal trend, researchers generally encouraged the continuous usage of MVS for food processing. Since ensemble and deep ensemble-based HSIC techniques as an AI models has demonstrated excellent attraction in classification performance, features including regeneration of sequential input data necessitates further investigation to complements the insufficient labeled data required from pixels to spectral band for HSI system efficiency across various realms. Fusion of deep ensembles with multimodal data (LiDAR or thermal imagery) also requires some exploration to provide comprehensive understanding of complicated scenes. This can be conducted by evaluation of high dimensionality, spectral variability and managing techniques that will help to overcome the difficulties of sparsity of training data accessibility (which is not limited to DA, AL, TL AEs and FSL). This is because, the aforementioned has demonstrated superb outcome in remote sensing technology and it is anticipated to tackle wastage of food along the production chain and help to contribute to food security.

## Declaration of interest statement

The authors declare that there is no any conflict of interest in the manuscript submitted for this publication.

## CRediT authorship contribution statement

**Olorunsogo Tunde Samuel:** Supervision. **Abdulwahab Ismail Durojaiye:** Supervision, Writting and editing the Manuscript. **Alkali Babawuya:** Supervision. **Bolanle Adenike Adejumo:** Supervision. **Ida Idayu Muhamad:** Supervision.

## Declaration of Competing Interest

We wish to officially declare our interest in publishing our review article with your reputable journal (Food and Humanity) as it has not been submitted to anywhere else for publication. We shall be highly delighted if the article is considered for publication. We also wish to indicate that there is no any conflict of interest amongst the authors.

## References


Abdel-Rahman, F., Okeremgbo, B., Alhamadah, F., Jamadar, S., Anthony, K., & Saleh, M. A. (2017). Caenorhabditis elegans as a model to study the impact of exposure to light emitting diode (LED) domestic lighting. *J. Environ. Sci. Health Part A Toxic/Hazard Subst. Environ. Eng., 52*, 433–439. https://doi.org/10.1080/10934529.2016.1270676

Adams, J. B., Sabol, D. E., Kapos, V., Almeida Filho, R., Roberts, D. A., Smith, M. O., & Gillespie, A. R. (1995). Classification of multispectral images based on fractions of endmembers: Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment, 52*(2), 137–154. https://doi.org/10.1016/0034-4257(94)00098-8

Aghbashlo, M., Hosseinpour, S., & Ghasemi-Varnamkhasti, M. (2014). Computer vision Food, technology for real-time food quality assurance during drying process. *Trends in Technology and Science, 39*. https://doi.org/10.1016/j.tifs.2014.06.003

Ahn, D. H., Choi, J. Y., Kim, H. C., Cho, J. S., Moon, K. D., & Park, T. (2019). Estimating the composition of food nutrients from hyperspectral signals based on deep neural networks. *Sensors (Switzerland), 19*(7). https://doi.org/10.3390/s19071560

Azhar, R., Tuwohingide, D., Kamudi, D., Sarimuddin, & Suciati, N. (2015). Batik image classification using SIFT feature extraction, bag of features and support vector machine. *Procedia Computer Science, 72*, 24–30. https://doi.org/10.1016/j.procs.2015.12.101

Bioucas-Dias, J. M., Plaza, A., Dobigeon, N., Parente, M., Du, Q., Gader, P., & Chanussot, J. (2012). Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*(2), 354–379. https://doi.org/10.1109/JSTARS.2012.2194696

Botero-Valencia, J. S., & Valencia-Aguirre, J. (2021). Portable low-cost IoT hyperspectral acquisition device for indoor/outdoor applications. *HardwareX, 10*, Article e00216. https://doi.org/10.1016/j.ohx.2021.e00216

Capitanio, F., Coppola, A., & Pascucci, S. (2010). Product and process innovation in the Industry., Italian food. *Agribusiness, 26*(4), 503–518. https://doi.org/10.1002/agr.20239

Cardello, A. V., Schutz, H. G., & Lesher, L. L. (2007). Consumer perceptions of foods processed Food, by innovative and emerging technologies: a conjoint analytic study. *Innovative Technology, Science Emergency, 8*(1), 73–83. https://doi.org/10.1016/j.ifset.2006.07.002

Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). "A simple framework for contrastive learning of visual representations,." Proceeding of the 37th International Conference on Machine Learning. Vienna, Austria, PMLR 119, 1597–1607. https://github.com/google-research/simclr.

Hong, D., Wu, X., Ghamisi, P., Chanussot, J., Yokoya, N., & Zhu, X. X. (2020). Invariant attribute profiles: A spatial-frequency joint feature extractor for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 58*(6), 3791–3808. https://doi.org/10.1109/TGRS.2019.2957251

Dalal, N., Triggs, B., Dalal, N., & Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection To cite this version: Histograms of Oriented Gradients for Human Detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 886–893. https://doi.org/10.1109/CVPR.2005.177.

Dixit, Y., Hitchman, S., Hicks, T. M., Lim, P., Wong, C. K., Holibar, L., Gordon, K. C., Loeffen, M., Farouk, M. M., Craigie, C. R., & Reis, M. M. (2021). Non-invasive spectroscopic and imaging systems for prediction of beef quality in a meat processing pilot plant. *Meat Science, 181*, Article 108410. https://doi.org/10.1016/j.meatsci.2020.108410

Drumetz, L., Henrot, S., Veganzones, M.A., Chanussot, J., & Jutten, C. (2015). Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability. *Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote Sensing* , 2015-June. https://doi.org/10.1109/WHISPERS.2015.8075417.

Eismann, M. T. (2012). *Hyperspectral Remote Sensing*. Bellingham, Washington DC: SPIE,. https://doi.org/10.1117/3.899758

ElMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for nondestructive determination of some quality attributes for strawberry. *Journal of Food Engineering, 81*(1), 98–107. https://doi.org/10.1016/j.jfoodeng.2006.10.016

ElMasry, G., Wang, N., ElSayed, A., & Ngadi, M. (2007). Hyperspectral imaging for non-destructive determination of some quality attributes for strawberry. *Journal of Food Engineering, 81*, 98–107.

Fan, S., Li, C., Huang, W., & Chen, L. (2017). Detection of blueberry internal bruising over Wavelengths, time using nir hyperspectral reflectance imaging with optimum. *Postharvest Biology and Technology, 134*, 55–66. https://doi.org/10.1016/j.postharvbio.2017.08.012

Feng, L., Wei, L., Nie, Y., Huang, M., Yang, L., Fu, X., & Zhou, J. (2019). Design of a compact spectrometer with large field of view based on freeform surface. *Opt. Commun, 444*, 81–86. https://doi.org/10.1016/j.optcom.2019.03.064

Fu, X., Ma, W. K., Bioucas-Dias, M., & Chan, T. H. (2016). Semiblind hyperspectral unmixing in the presence of spectral library mismatches. *IEEE Trans. Geosci. Remote Sens., 54*(9), 5171–5184. https://doi.org/10.1109/TGRS.2016.2557340

Garcia, R., Nicosevici, T., & Cufi, X. (2003). *On the way to solve lighting problems in underwater imaging* (pp. 1018–1024). IEEE Xplore, 2,. https://doi.org/10.1109/OCEANS.2002.1192107

Gholian-Jouybari, F., Hashemi-Amiri, O., Mosallanezhad, B., & HajiaghaeiKeshteli, M. (2023). Metaheuristic algorithms for a sustainable agri-food supply chain considering marketing practices under uncertainty. *Expert Syst. Appl., 213*, Article 118880. https://doi.org/10.1016/J.ESWA.2022.118880

Grigoriev, N. (2022). Construction and development of a low-cost hyperspectral imaging system. Master's Thesis in Engineering Physics, Umeå University, Faculty of Science and Technology, Department of Physics. pp. 1- 18. (Accessed on 28th, January, 2024).

Gómez Manzanares, Á., Vázquez Moliní, D., Alvarez Fernandez-Balbuena, A., Mayorga Pinilla, S., & Martínez Antón, J. C. (2022). Measuring High Dynamic Range Spectral Reflectance of Artworks through an Image Capture Matrix Hyperspectral Camera. *Sensors, 22*, 4664. https://doi.org/10.3390/s22134664

Habel, R., Kudenov, M., & Wimmer, M. (2012). Practical spectral photography. *Computer Graphics Forum, 31*(2), 449–458. https://doi.org/10.1111/j.1467-8659.2012.03024.x

Haut, J. M., Paoletti, M. E., Plaza, J., Plaza, A., & Li, J. (2019). Hyperspectral image classification using random occlusion data augmentation. *IEEE Geoscience and Remote Sensing Letters, 16*(11), 1751–1755. https://doi.org/10.1109/LGRS.2019.2909495

He, K. M., Zhang, X. Y., Ren, S. Q., & Sun, J. (2016). Deep residual learning for image recognition. In. *IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. https://doi.org/10.1109/CVPR.2016.90

He, H.-J., Wang, Y., Ou, X., Ma, H., Liu, H., & Yan, J. (2023). Rapid determination of chemical compositions in chicken flesh by mining hyperspectral data. *Journal of Food Composition and Analysis, 116*, Article 105069. https://doi.org/10.1016/j.jfca.2022.105069

Hecht-Nielsen, R. (1992). Theory of the backpropagation neural network. In: Neural Perception., Networks for Perception. *Elsevier*, 65–93. https://doi.org/10.1016/B978–0-12–741252-8.50010–8.

Henriksen, M. B., Prentice, E. F., van Hazendonk, C. M., Sigernes, F., & Johansen, T. A. (2022). Do-it-yourself VIS/NIR pushbroom imager hyperspectral with C-mount optics. *Optics Continuum, 1*(2), 427. https://doi.org/10.1364/optcon.450693

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Hong, D., Member, S., & Yokoya, N. (2019). An augmented linear mixing model to address spectral variability for hyperspectral unmixing. *IEEE Transactions on Image Processing, 28*(4), 1923–1938. https://doi.org/10.1109/TIP.2018.2878958

Hong, D., Zhang, B., Li, X., Li, Y., Li, C., Yao, J., Yokoya, N., Li, H., Ghamisi, P., Jia, X., Plaza, A., Gamba, P., Benediktsson, J. A., & Chanussot, J. (2024). SpectralGPT: spectral remote sensing foundation model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1–16. https://doi.org/10.1109/TPAMI.2024.3362475

Høye, G., Løke, T., & Fridman, A. (2015). Method for quantifying image quality in push-broom hyperspectral cameras. *Optical Engineering, 54*(5). https://doi.org/10.1117/1.OE.54.5.053102

Hsueh, M. H., Lai, C. J., Wang, S. H., Zeng, Y. S., Hsieh, C. H., Pan, C. Y., & Huang, W. C. (2021). Effect of printing parameters on the thermal and mechanical properties of 3d-printed pla and petg, using fused deposition modeling. *Polymers, 13*(11). https://doi.org/10.3390/polym13111758

Hu, M., Dong, Q., & Liu, B. L. (2016). Classification and characterization of blueberry mechanical damage with time evolution using reflectance, transmittance and Spectroscopy, interactance imaging. *Comput Electron Agric, 122*, 19–28. https://doi.org/10.1016/j.compag.2016.01.015

Hu, X., Li, T., Zhou, T., & Peng, Y. (2021). Deep spatial–spectral subspace clustering for hyperspectral images based on contrastive learning. *Remote Sens, 13*(21), 4418. https://doi.org/10.3390/rs13214418

Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing, 7*(11), 14680–14707. https://doi.org/10.3390/rs71114680

Huang, L., Chen, C., Li, W., & Du, Q. (2016). Remote sensing image scene classification using multi-scale completed local binary patterns and fisher vectors. *Remote Sensing, 8*(6), 1–17. https://doi.org/10.3390/rs8060483

Huang, X., Dong, M., Li, J., & Guo, X. (2022). A 3-D-Swin transformer-based hierarchical contrastive learning method for hyperspectral image classification. Art. no. 5411415 *IEEE Trans. Geosci. Remote Sens., 60*. https://doi.org/10.1109/TGRS.2022.3202036.

Hussain, A., Pu, H., & Sun, D. W. (2018). Innovative nondestructive imaging techniques for ripening and maturity of fruits–a review of recent applications. *Trends in Food Science & Technology, 72*, 144–152. https://doi.org/10.1016/j.tifs.2017.12.010

Jia, B., Wang, W., Ni, X., Lawrence, K. C., Zhuang, H., Yoon, S. C., & Gao, Z. (2020). Essential processing methods of hyperspectral images of agricultural and food products. *In Chemometrics and Intelligent Laboratory Systems* (Vol. 198). https://doi.org/10.1016/j.chemolab.2020.103936

Jiang, T., Li, Y., Xie, W., & Du, Q. (2020). Discriminative reconstruction constrained generative adversarial network for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens., 58*(7), 4666–4679.

Kabir, M. H., Guindo, M. L., Chen, R., Sanaeifar, A., & Liu, F. (2022). Application of Laser-Induced Breakdown Spectroscopy and Chemometrics for the Quality Evaluation of Foods with Medicinal Properties: A Review. *Foods, 11*(14), 2–19. https://doi.org/10.3390/foods11142051

Jiang, H., Diao, Z., Shi, T., Zhou, Y., Wang, F., Hu, W., & Yao, Y. D. (2023). A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation. *Computers in Biology and Medicine, 157*, Article 106726. https://doi.org/https://doi.org/10.1016/J. COMPBIOMED.2023.106726.

Kieu, B. Y., Jensen, C. S., & Jensen. (2018). Outlier detection for multidimensional time series using deep neural networks. In *Proc. 19th IEEE Int. Conf. Mobile Data Manage (MDM)* (pp. 125–134).

Khoshtaghaza, M. H., Khojastehnazhand, M., Mojaradi, B., Goodarzi, M., & Saeys, W. (2016). Texture quality analysis of rainbow trout using hyperspectral imaging method. *Int. J. Food Prop., 19*, 974–983. https://doi.org/10.1080/10942912.2015.1042111

Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. *In Proceedings of the 25th International Conference on Neural Information Processing Systems*, 1097–1105. https://doi.org/10.1145/3065386.

Langelaan, H., Pereira da Silva, F., Thoden van Velzen, U., Broeze, J., Matser, A., & Vollebregt, M., Schroϵen, K. (2013). Technology options for feeding 10 billion and, In: Options for sustainable food processing. State of the art report. Science Technology Options Assessment. Brussels, European Parliament. Http://Www.Europarl.Europa.Eu/RegData/Etudes/Etudes/Join/2013/5135 33/IPOLJOIN_ET (2013)513533_EN.Pdf.

Li, C., Zhang, B., Hong, D., Yao, J., & Chanussot, J. (2023a). LRR-Net: An interpretable deep unfolding network for hyperspectral anomaly detection. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–12. https://doi.org/10.1109/TGRS.2023.3279834

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. *2021 IEEE/CVF International Conference on Computer Vision (ICCV)* , 9992–10002. https://doi.org/10.1109/ICCV48922.2021.00986.

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic and, segmentation. In: *Proceedings of the IEEE Conference on Computer Vision Pattern Recognition* (CVPR), Boston, 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965.

Lowe, D.G. (1999). "Object recognition from local scale-invariant features,." *In Proceedings of the Seventh IEEE International Conference on Computer Vision*, *2*, 1150–1157. https://doi.org/10.1109/ICCV.1999.790410.

Li, J., Li, X., & Yan, Y. (2023b). Unlocking the potential of data augmentation in contrastive learning for hyperspectral image classification. *Remote Sensing, 15*(12), 3123. https://doi.org/10.3390/rs15123123

Lu, X., Zhang, W., & Huang, J. (2020). Exploiting embedding manifold of autoencoders for hyperspectral anomaly detection. *IEEE Trans. Geosci. Remote Sens., 58*(3), 1527–1537. https://doi.org/10.1109/TGRS.2019.2944419

Lyu, H., & Lu, H. (2016). Learning a transferable change detection method by recurrent neural network. In *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)* (pp. 5157–5160). IEEE.

Ma, X., & McSweeney, P. (2008). Product and process innovation in the food processing Review, industry: case study in guangxi province. *Australasian Agribusiness, 16* (1673), 2016-136765. ISSN 1442-6951.

Mahendran, R., Palanivel, J., & Varadarajan, E. (2021). Influence of Processing Parameters and PYTHON Based Image Analysis for Quantification of Carcinogenic Acrylamide in Potato Chips. *Chemistry Africa, 4*(3), 669–675. https://doi.org/10.1007/s42250-021-00237-9

Manzini, R., & Accorsi, R. (2013). The new conceptual framework for food supply chain assessment. *J. Food Eng., 115*(2), 251–263. https://doi.org/10.1016/j.jfoodeng.2012.10.026

Mao, Y., Betters, C. H., Evans, B., Artlett, C. P., Leon-Saval, S. G., Garske, S., Cairns, I. H., Cocks, T., Winter, R., & Dell, T. (2022). OpenHSI: A Complete Open-Source Hyperspectral Imaging Solution for Everyone. *Remote Sensing, 14*(9), 1–19. https://doi.org/10.3390/rs14092244

Myers, A., Johnston, N., Rathod, V., Korattikara, A., Gorban, A., Silberman, N., Guadarrama, S., Papandreou, G., Huang, J., & Murphy, K. (2015). Im2Calories: Towards an Automated Mobile Vision Food Diary. *2015 IEEE International Conference on Computer Vision (ICCV)* , 1233–1241. https://doi.org/10.1109/ICCV.2015.146.

Nhat, H.T. M., & Hoang, V.T. (2019). Feature fusion by using LBP, HOG, GIST descriptors and Canonical Correlation Analysis for face recognition. 26th International Conference on Telecommunications (ICT), 371–375). IEEE. https://doi.org/10.1109/ICT.2019.8798816.

Norris, D. (2020). Machine Learning with the Raspberry Pi: Experiments with Data and Computer Vision. *Springer*, *Berlin/Hei*. Pihttps://doi.org/10.1007/978–1-4842–5174-4_1.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*(3), 145–175. https://doi.org/10.1023/A:1011139631724

Othman, S., Mavani, N. R., Hussain, M. A., Rahman, N. A., & Mohd Ali, J. (2023). Artificial intelligence-based techniques for adulteration and defect detections in food and agricultural industry: A review. *Journal of Agriculture and Food Research, 12* (April), Article 100590. https://doi.org/10.1016/j.jafr.2023.100590

Palmason, J.A., Benediktsson, J.A., & Sveinsson, J.R. (2005). Classification of hyperspectral ROSIS data from urban areas. *RAST 2005 - Proceedings of 2nd International Conference on Recent Advances in Space Technologies, 2005*(3), 63–69. https://doi.org/10.1109/RAST.2005.1512536.

Pechlivani, E. M., Papadimitriou, A., Pemas, S., Giakoumoglou, N., & Tzovaras, D. (2023). Low-cost hyperspectral imaging device for portable remote sensing. *Instruments, 7*(4), 32. https://doi.org/10.3390/instruments7040032

Prentice, E. F., Grøtte, M. E., Sigernes, F., & Johansen, T. A. (2021). Design of a hyperspectral imager using COTS optics for small satellite applications. *1*, 187. https://doi.org/10.1117/12.2599937

Riihiaho, K. A., Eskelinen, M. A., & Pölönen, I. (2021). A Do-It-Yourself hyperspectral imager brought to practice with open-source python. *Sensors, 21*(4), 1–17. https://doi.org/10.3390/s21041072, 1072.

Rogge, D. M., Rivard, B., Jinkai, Z., & Jilu, F. (2006). Iterative spectral unmixing for optimizing per-pixel endmember sets. *IEEE Transactions on Geoscience and Remote Sensing, 44*(12), 3725–3735. https://doi.org/10.1109/TGRS.2006.881123

Roy, S. K., Hong, D., Kar, P., Wu, X., Liu, X., & Zhao, D. (2021). Lightweight heterogeneous kernel convolution for hyperspectral image classification with noisy labels. *IEEE Geoscience and Remote Sensing Letters, 99*, 1–5. https://doi.org/10.1109/LGRS.2021.3112755

Roy, S. K., Manna, S., Dubey, S. R., & Chaudhuri, B. B. (2023). LiSHT: Non-parametric Linearly Scaled Hyperbolic Tangent Activation Function for Neural Networks. *Communications in Computer and Information Science, 1776 CCIS*, 462–476. https://doi.org/10.1007/978-3-031-31407-0_35

Saha, D., & Manickavasagan, A. (2021). Current research in food science machine learning techniques for analysis of hyperspectral images to determine quality of food products: A review. *Current Research in Food Science, 4*, 28–44. https://doi.org/10.1016/j.crfs.2021.01.002

Salazar-Vazquez, J., & Mendez-Vazquez, A. (2020). A plug-and-play Hyperspectral Imaging Sensor using low-cost equipment. *HardwareX, 7*, Article e00087. https://doi.org/10.1016/j.ohx.2019.e00087

Shafiee, S., Minaei, S., Moghaddam-Charkari, N., & Barzegar, M. (2014). Honey characterization using computer vision system and artificial neural networks. *Food Chemistry, 159*, 143–150. https://doi.org/10.1016/j.foodchem.2014.02.136

Sigernes, F., Syrjäsuo, M., Storvold, R., Fortuna, J., Grøtte, M. E., & Johansen, T. A. (2018). Do it yourself hyperspectral imager for handheld to airborne operations. *Optics Express, 26*(5), 6021. https://doi.org/10.1364/oe.26.006021

Song, W., Li, S., Fang, L., & Lu, T. (2018). Hyperspectral image classification with deep feature fusion network. *IEEE transactions on geoscience and remote sensing., 56*(6), 3173–3184.

Skauli, T. (2017). "Feasibility of a standard for full specification of spectral imager performance," in Hyperspectral Imaging Sensors: *Innovative Applications and Sensor Standards*, *Bannon*, *D.* https://doi.org/10.1117/12.2262785.

Sun, Y., Xiao, H., Tu, S., Sun, K., Pan, L. Q., & Tu, K. (2017). Detecting decay pearch using a rotating hyperspectral imaging testbed. *LWT Food Science and Technology, 87*. https://doi.org/10.1016/j.lwt.2017.08.086

Szegedy, C., Liu, W., Jia, Y. Q., Sermanet, P., Reed, S., Anguelov, D., & Rabinovich, A. (2015). Going deeper with convolutions. In. *5*. *IEEE Conference on Computer Vision and Pattern Recognition*, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Thomas, S., Kuska, M. T., Bohnenkamp, D., Brugger, A., Alisaac, E., Wahabzada, M., Behmann, J., & Mahlein, A. K. (2018). Benefits of hyperspectral imaging for plant disease detection and plant protection: A technical perspective. *J. Plant Dis. Prot., 125*(1), 5–20. https://doi.org/10.1007/s41348-017-0124-6

Thouvenin, P. A., Dobigeon, N., & Tourneret, J. Y. (2016). Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. *IEEE Transactions on Signal Processing, 64*(2), 525–538. https://doi.org/10.1109/TSP.2015.2486746

Ullah, F., Ullah, I., Khan, R. U., Khan, S., Khan, K., & Pau, G. (2024). Conventional to deep ensemble methods for hyperspectral image classification: A comprehensive survey. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17*, 3878–3916. https://doi.org/10.1109/JSTARS.2024.3353551

Uto, H., Seki, K., Saito, G., & Kosugi, Y. (2015). Development of a low-cost, lightweight hyperspectral imaging system based on a polygon mirror and compact spectrometers. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 9*(2), 861–875.

Uto, K., Seki, H., Saito, G., & Kosugi, Y. (2016). Development of a low-cost hyperspectral whiskbroom imager using an optical fiber bundle, a swing mirror, and compact spectrometers. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens., 9*(9), 3909–3925.

Van Der Goot, A. J., Pelgrom, P. J., Berghout, J. A., Geerts, M. E., Jankowiak, L., Hardt, N. A., Keijer, J., Schutyser, M. A., Nikiforidis, C. V., & Boom, R. M. (2016). Concepts for further sustainable production of foods. *Journal of Food Engineering, 168*, 42–51. https://doi.org/10.1016/j.jfoodeng.2015.07.010

Veganzones, M. A., Drumetz, L., Tochon, G., Dalla Mura, M., Plaza, A., Bioucas-Dias, J., & Chanussot, J. (2014). A new extended linear mixing model to address spectral variability. *Workshop on Hyperspectral Image and Signal Processing. 2014-June Evolution in Remote Sensing*. https://doi.org/10.1109/WHISPERS.2014.8077595.

Wan, P., Toudeshki, A., Tan, H., & Ehsani, R. (2018). A methodology for fresh tomato Agric., maturity detection using computer vision. *Computer and Electronics in Agriculture, 146*, 43–50. https://doi.org/10.1016/j.compag.2018.01.011

Wang, W., Heitschmidt, G. W., Ni, X., Windham, W. R., Hawkins, S., & Chu, X. (2014). Identification of aflatoxin b1, on maize kernel surfaces using hyperspectral imaging. *Food Control, 42*, 78–86. https://doi.org/10.1016/j.foodcont.2014.01.038

Wang, Z., Hu, M., & Zhai, G. (2018). Application of deep learning architectures for accurate and rapid detection of internal mechanical damage of blueberry using hyperspectral transmittance data. *Sensors, 18*, 1126.

Wang, Y., Song, T., Xie, Y., & Roy, S. K. (2022). y, "A probabilistic neighbourhood pooling-based attention network for hyperspectral image classification. *Remote Sensing Letters, 13*(1), 65–75. https://doi.org/10.1080/2150704X.2021.1992034

Williams, P. J., Geladi, P., Britz, T. J., & Manley, M. (2012). Investigation of fungal development in maize kernels using NIR hyperspectral imaging and multivariate data analysis. *J. Cereal Sci., 55*, 272–278. https://doi.org/10.1016/j.jcs.2011.12.003

Wu, D., Sun, D. W., & He, Y. (2012). Application of long-wave near infrared hyperspectral imaging for measurement of color distribution in salmon fillet. *Innovative Food Science & Emerging Technologies, 16*, 361–372. https://doi.org/10.1016/j.ifset.2012.08.003

Wu, X., Hong, D., & Chanussot, J. (2023). UIU-Net: U-Net in U-Net for infrared small object detection. *" Trans. Image Process,, 32*, 364–376. https://doi.org/10.1109/TIP.2022.3228497

Xie, W., Zhang, J., Lei, J., Li, Y., & Jia, X. (2021). Self-spectral learning with GAN based spectral–spatial target detection for hyperspectral image. *Neural Network, 142*, 375–387.

Yang, Q., Shi, L., Han, J., Zha, Y., & Zhu, P. (2019). Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res, 235*, 142–153.

Yao, H., Chen, R., Chen, W., Sun, H., Xie, W., & Lu, X. (2023). Pseudolabel-based unreliable sample learning for semi-supervised hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–16. https://doi.org/10.1109/TGRS.2023.3322558

Yoon, S. C., Lawrence, K. C., Line, J. E., Siragusa, G. R., Feldner, P. W., Park, B., & Windham, W. R. (2010). Detection of campylobacter, colonies using hyperspectral imaging. *Sens. Instrumen. Food Qual. Saf., 4*, 35–49. https://doi.org/10.1007/s11694-010-9094-0

Zareiforoush, H., Minaei, S., Alizadeh, M. R., & Banakar, A. (2015). Qualitative classification of milled rice grains using computer vision and metaheuristic techniques. *Journal of Food Science and Technology, 53*(1), 118–131. https://doi.org/10.1007/s13197-015-1947-4

Zareiforoush, H., Minaei, S., Alizadeh, M. R., Banakar, A., & Samani, B. H. (2016). Design Rice, development and performance evaluation of an automatic control system for Electron., whitening machine based on computer vision and fuzzy logic. *Comput Electron Agric., 124*, 14–22. https://doi.org/10.1016/j.compag.2016.01.024

Zhou, L., Zhang, C., Liu, F., Qiu, Z., & He, Y. (2019). Application of deep learning in food: a review. *Comprehensive Reviews in Food Science and Food Safety,. Blackwell Publishing Inc, 18*(6), 1793–1811. https://doi.org/https://doi.org/10.1111/ 1541-4337.12492.

Zhu, L., Spachos, P., Pensini, E., & Plataniotis, K. N. (2021). Deep learning and machine vision for food processing: A survey. *Curr. Res. Food Sci, 4*, 233–249. https://doi.org/10.1016/j.crfs.2021.03.009

Zhu, L., & Spachos, P. (2020). Food grading system using support vector machine and YOLOv3 methods. Proceedings - *IEEE Symposium on Computers and Communications* , 2020-July. https://doi.org/10.1109/ISCC50000.2020.9219589.

Zurada, J.M. (1992). *Introduction to Artificial Neural Systems,. 8*, (West St. Paul). West Publishing Co., United States. 683 ISBN: 0314933913.