Indonesian Journal of Science & Technology

Journal homepage: http://ejournal.upi.edu/index.php/ijost/

Enhanced Product Defect Detection in Smart Manufacturing Using ConvNeXt-Stacked Autoencoder Architecture

Ahmed M. Hasan[1], Ahmed R. Nasser[1], Ahmed Sabah Al-Araji[1], Dalia Abdulkareem Shafiq[2], Abdulkareem Sh. Mahdi Al-Obaidi[2], Huthaifa Al-Khazraji[1], Amjad J. Humaidi[1,*]

[1] University of Technology-Iraq, Baghdad, Iraq
[2] Taylor's University, Subang Jaya, Malaysia
*Correspondence: E-mail: amjad.j.humaidi@uotechnology.edu.iq

A B S T R A C T

In the era of Industry 4.0, ensuring product quality through accurate and efficient defect detection has become essential because traditional manual inspection methods are often time-consuming and prone to inconsistency. This study aims to enhance defect detection in smart manufacturing by proposing a hybrid deep learning architecture that combines ConvNeXt and stacked autoencoders. The method leverages ConvNeXt for robust feature extraction and stacked autoencoders for efficient data reconstruction and classification. The model was trained and tested on real-world industrial image datasets involving damaged and intact packaging. Results demonstrate that the proposed method outperforms conventional convolutional neural networks in both detection accuracy and processing efficiency because of its ability to extract deep spatial and semantic features. This research contributes to the advancement of autonomous quality control systems in smart manufacturing. Its impact lies in reducing human dependency, improving inspection accuracy, and fostering the development of intelligent, self-regulating production lines.

1. INTRODUCTION

Advanced technologies such as information technology (IT), Internet of Things (IoT), artificial intelligence (AI), augmented reality (AR), virtual reality (VR), sensor networks, computerized controls, and big data drive manufacturing companies to work in a manner to satisfy the necessities of automatic manufacturing systems [1]. Many reports regarding this matter have been well-developed (**Table 1**).

Table 1. Previous studies on advanced technology to support industry 4.0.

No	Topic	Title	Ref.
1.	Information Technology (IT)	Theoretical foundations of the creation of a curriculum in higher project IT in education	[2]
2.		Improving training of modern leaders utilizing IT in the administration of the higher education system	[3]
3.		The paradigm of curriculum differentiation in higher IT education	[4]
4.		Smart city and society 5.0: Involvement of information technology in the development of public service systems in Indonesia	[5]
5.		Internet of things-based child stunting detection system for supporting sustainable development goals	[6]
6.		Predictive deep learning models to identify traumatic brain injuries using MRI data	[7]
7.		Enhancing digital literacy and teacher-preneurship through a critical pedagogy-based training platform	[8]
8.		Electric vehicle consumption dataset tailored to Malaysian situation and implemented using rapid miner auto-model	[9]
9.		Emerging applications of iot, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis	[10]
10.	Internet of Thing (IoT)	Mesh network based on MQTT broker for smart home and iiot factory	[11]
11.		Easy-mushroom mobile application using the Internet of Things (iot)	[12]
12.		Greening the internet of things: A comprehensive review of sustainable IOT solutions from an educational perspective	[13]
13.		Internet of things-based child stunting detection system for supporting sustainable development goals	[14]
14.		Emerging applications of iot, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis	[15]
15.		Emerging trends in technology: insights across machine learning, digitalization, and industry applications	[16]
16.		Internet of things for monitoring and optimisation of stand-alone systems in rural area: an experimental case	[17]
17.		Mobiledfu: classification of diabetic foot ulcer infection on the edge	[18]
18.	Artificial Intelligence (AI)	Chatbot artificial intelligence as educational tools in science and engineering education: A literature review and bibliometric mapping analysis with its advantages and disadvantages	[19]
19.		How bibliometric analysis using vosviewer based on artificial intelligence data (using researchrabbit Data): Explore research trends in hydrology content	[20]
20.		Artificial intelligence (AI)-based learning media: Definition, bibliometric, classification, and issues for enhancing creative thinking in education	[21]
21.		Trends in the use of artificial intelligence (AI) technology in increasing physical activity	[22]

Table 1 (continue). Previous studies on advanced technology to support industry 4.0.

No	Topic	Title	Ref.
22.		Bibliometric analysis of research trends in conceptual understanding and sustainability awareness through artificial intelligence (AI) and digital learning media	[23]
23.		The future of learning: ethical and philosophical implications of artificial intelligence (AI) integration in education	[24]
24.		University students' awareness of, access to, and use of artificial intelligence for learning in Kwara State	[25]
25.		Bibliometric analysis on artificial intelligence research in Indonesia vocational education	[26]
26.		Primary education undergraduates' competency in the use of artificial intelligence for learning in Kwara State	[27]
27.		The digital frontier: AI-enabled transformations in higher education management	[28]
28.		The role of chatgpt AI in student learning experience	[29]
29.	Sensor	Monitoring of air quality with satellite-based sensor: The case of four towns in Southeast, Nigeria	[30]
30.		Vehicle-tracking mobile application without a GPS sensor	[31]
31.		Securing wireless sensor networks, types of attacks, and detection/prevention techniques: An educational perspective	[32]
32.		Graphene-based electrochemical sensors for heavy metal ions detection: A comprehensive review	[33]
33.		Internet of things-based child stunting detection system for supporting sustainable development goals	[34]
34.		Emerging trends in technology: insights across machine learning, digitalization, and industry applications	[35]
35.		Interference mitigation for dynamic user connectivity using sdn and radio resource management in cell-less networks	[36]
36.		Internet of things for monitoring and optimisation of stand-alone systems in rural area: an experimental case	[36]
37.		Mobiledfu: classification of diabetic foot ulcer infection on the edge	[38]
38.		Development of augmented reality application for exercise to promote health among elderly	[39]
39.		Application of augmented reality technology with the fuzzy logic method as an online physical education lecture method in the new normal era	[40]
40.		How to create augmented reality (AR) applications using unity and vuforia engine to teach basic algorithm concepts: Step-by-step procedure and bibliometric analysis	[41]
41.		Emerging applications of iot, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis	[42]
42.		Emerging trends in technology: insights across machine learning, digitalization, and industry applications	[43]
43.		Let's immerse in metaverse: overview, challenges, and stumeta framework for successful implementation	[44]
44.		Development of science virtual laboratory (scivlab) to develop critical thinking skills in elementary schools on the topic of changes in the state of substances	[45]
45.		Usability assessment of flipo-ar: navigating learning in a vuca world with augmented reality	[46]
46.	Virtual Reality (VR)	The use of virtual reality as a substitute for the pre-school students' field trip activity during the learning from home period	[47]

Table 1 (continue). Previous studies on advanced technology to support industry 4.0.

No	Topic	Title	Ref.
47.		Colleges of education lecturers' attitude towards the use of virtual classrooms for instruction	[48]
48.		Students' learning experiences and preference in performing science experiments using hands-on and virtual laboratory	[49]
49.		The effectiveness of using a virtual laboratory in distance learning on the measurement materials of the natural sciences of physics for junior high school students	[50]
50.		Perception of early childhood education lecturers on the use of virtual learning	[51]
51.		Lecturers perceived proficiency in the use of virtual classrooms for instruction in colleges of education	[52]
52.		Development and acceptability of virtual laboratory in learning systematics	[53]
53.		Utilization of virtual reality chat as a means of learning communication in the field of education	[54]
54.		The effectiveness of using a virtual laboratory in distance learning on the measurement materials of the natural sciences of physics for junior high school students.	[55]
55.		Developing an identification and testing system for cardiac surgical instrument capabilities	[56]
56.		Emerging applications of iot, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis	[57]
57.		Emerging trends in technology: insights across machine learning, digitalization, and industry applications	[58]
58.		Let's immerse in metaverse: overview, challenges, and stumeta framework for successful implementation	[59]
59.		Development of science virtual laboratory (scivlab) to develop critical thinking skills in elementary schools on the topic of changes in the state of substances	[60]

With the appearance of the Fourth Industrial Revolution (Industry 4.0), the concept of product quality control has gained high attention in industrial manufacturing. One of the important research directions in the field of quality control, improving the detection of product defects during the manufacturing process. Delivering products without defects is always of major concern to top decision-makers in production companies. Faults and imperfections like internal holes, abrasions, and also scratches might occur in the production of products. In such cases, the quality of products will affect production efficiency, since products with poor quality considered a waste of raw materials and this is costs [61]. Automatic defect-detection systems utilizing advanced technologies such as advanced sensors and artificial intelligence have noticeable advantages over the time-consuming traditional manual detection [62]. Besides, the performance of human inspection gets very exhausted with repetitive tasks and hence these tasks are typically very labor-intensive. However, automatic detection of effective visual defects in the product, which intends to identify a possible defective area of a product image used and then categorize these images into defect and defect-free, appears as a solution to the problem [63].

Artificial Intelligence (AI), in particular Machine Learning (ML) as well as Deep Learning (DL), can both be defined as algorithms consisting of multiple processing layers and can learn from data. In other words, algorithms take input data, train themselves to observe patterns

found in the data, and afterward predict the output for a new set of data [64-66]. In terms of support in decision-making, ML and DL algorithms are revealing massive impending in the scrutiny of large amounts of data, currently readily available, aiming to enhance the efficiency of the proposed algorithm [67]. With the significant enhancement of the Graphical Processing Units (GPU) computing aptitudes, DL approaches are mostly known to have remarkable rewards, in terms of the capability to conduct multivariate, high dimensional data and hence can extract concealed relationships within the specified data. Therefore, these algorithms appeared to be one of the most dominant tools in several applications and wide field of research [68-69].

Various research and studies have been performed on product defect detection in different industrial sectors such as semiconductor, carpet, steel, and fabric manufacturing based on DL algorithms. For instance, [63] presented an application of Deep Neural Network (DNN) collective with a high-resolution optical quality camera to boost the precision of an industrial visual assessment in the printing process. To show the role of DL algorithms in enabling production companies to transfer to smart manufacturing, [70] proposed a visual quality control system using DL methods. A camera was placed over the production line. Then, the image of the product is sent to the algorithm to decide whether the product is "okay" or "not okay". To improve the productivity of the textile industry, [71] proposed an EfficientDet-D0 algorithm for a fabric defect detection system. Besides, [72] proposed a one-class classification (OCC) for carpet defect detection systems. The proposed system is trained only with normal samples whereas, throughout the test phase, both normal and defective images are employed. Three models are used. Convolutional autoencoder is first used as a hidden feature extractor. Therefore, the extorted feature vectors are then fed into the dimensionality reduction process by utilizing the method of principal component analysis (PCA). Finally, the training of the one-class classifier called Support Vector Data Description (SVDD) is performed using the above-mentioned resulting reduced-dimensional data.

All these studies show the capability of DL algorithms to settle independently about product quality devoid of human participation. However, improving the efficiency and exploring the capability of other DL algorithms for better performance has been an ongoing endeavor. Due to its higher accuracy and balanced precision and recall metrics as compared to other DL algorithms, in this paper, an improvement model of ConvNeXt with Stacked Autoencoders (ConvNeXt-SAEs) is proposed for product defect detection systems.

2. DEFECTS DETECTION SYSTEM

Most recently defect-detection system is an interesting topic in many fields, among them industries. It is widely used in production companies to ensure the quality of the product. Defects such as internal holes, abrasions, and even scratches might occur in the production of products. Therefore, a product defect detection system refers to the detection tools that are used to identify the external (i.e. surface) and/or internal defects of products [73]. Recently, automated defect-detection system utilizing integrated technologies such as image processing, pattern recognition, and DL algorithms has a broadly applied over manual detection to improve the performance of defect-detection systems [74]. A product defect detection system based on DL algorithms is built upon the data that is collected from the production process. Then, by extracting patterns from the data, DL algorithms can be used to detect whether the product is "okay" or "not okay". This process works as a decision-making for quality-enhancing measures [75]. Therefore, using an automated defect detection system can reduce costs and improve the efficiency of the production process. It also is considered one of the requirements for the transformation to the smart manufacturing industry [62]. As

one of the important research directions in the field of improving the performance of the automated defect detection system, is improving the software and algorithms in image processing. In this direction, a significant amount of interest has been increased in proposing, exploring, and evaluating modern DL architectures.

3. METHODS

To satisfy the requirement of the transition to smart manufacturing systems, we proposed a DL algorithm named ConvNeXt Stacked Autoencoder for product defect detection systems. In the literature, the performance of the ConvNeXt and the Stacked Autoencoders networks has been examined separately. In this research, the advantages of the two methods are combined and applied to the defect-detection system.

3.1. ConvNeXt

Self-attention-based deep learning methods are approaches that have many application areas such as image and language processing, and and temporal prediction models. In these approaches, the inputs focus on calculating the relationships between different regions. Self-attention-based methods aim to emphasize the features that are significant for the target outputs. These methods consider self-attention matrices as the main components. Relationship calculations are performed with these matrices and the relationship calculations are associated with the weights in the feature maps [76]. The self-attention-based method that can process images is the ConvNeXt model [77]. ConvNeXt is a pure convolutional neural network (CNN) built entirely from standard convolutional modules combined with principles of transformer network (i.e. vision transformer and the swing transformer) [78] The motivation behind ConvNeXt was to reexamine the design space of CNN network architecture and investigate the restrictions of what an untainted convolutional network can achieve. The researchers progressively modernized a standard ResNet towards an image transformer design and discovered numerous key components that contributed positively to the performance difference along the system. Therefore, the result of this discovery is a kind of pure convolutional model called ConvNeXt [79]. Built totally from standard convolutional modules, ConvNeXt competes positively with transformer-based models in COCO detection along with ADE20K segmentation while maintaining the simplicity and efficiency of standard convolutional networks.

The ConvNeXt model has approximately the same number of parameters and memory usage rate as the reference methods, but certain modules have been simplified with this model. The ConvNeXt model, whose layer principle is shown in **Figure 1**, first processes the input with a convolution layer. Then, a linear normalization process is applied to the features created by these convolution operations. With linear normalization, the convolution layer outputs are brought to a certain range. The last function of the ConvNeXt block is the activation of the Gaussian Error Linear Unit (GeLU) [80]. This activation process compresses the outputs and completes the transformation process. The layers and mathematical formulations of ConvNeXt are described as follows.
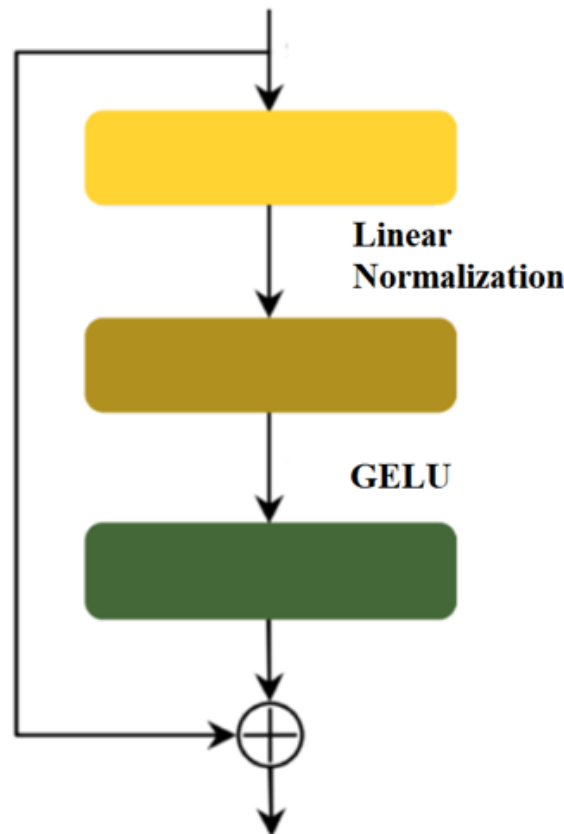
Figure 1. ConvNeXt block diagram.

(i)   Convolution: ConvNeXt relies on the basic convolution operation for image processing, where each output is computed as a weighted sum of neighboring points (see equation (1)):

$$Yi_{,j} = \sum x \cdot w \tag{1}$$

(ii)   Layer Normalization (LN): Layer normalization is used to stabilize the training process by normalizing the data across different feature channels (see equation (2)):

$$LN(x) = \frac{x - u}{\sqrt[2]{\sigma^2 - \epsilon}} \cdot \gamma - \beta \tag{2}$$

where u, β, γ and σ are the values learned during training.

(iii)   Deepwise Convolution: In this process, a separate convolution is applied to each channel, which reduces the computational volume (see equation (3)):

$$y_{i,j}^c = \sum x . w^c \tag{3}$$

where c denotes the channel number.

(iv)   Pointwise Convolution: A 1x1 convolution that combines information from different channels (see equation (4)):

$$y_{i,j} = \sum x^c . w^c \tag{4}$$

(v)   Stochastic Depth: During training, some remaining connections are randomly skipped with probability p. Equation (5) is in the following:

$$y = \begin{cases} Resblock(x) + x \ with \ p \\ x \ with \ (1 - p) \end{cases} \tag{5}$$

(vi)   GELU activation function: It is a smooth activation function that is better than ReLU (see equation (6)):

$$GELU(x)=x \cdot \Phi(x) \tag{6}$$

where $\Phi(x)$ is the cumulative distribution function of Gauss.

ConvNeXt combines these operations to enhance efficiency and performance, by blending traditional neural network techniques with some approaches inspired by the modern Transformers architecture.

3.2. Stacked Autoencoders

Stacked autoencoders (SAEs) are one of the artificial neural network models that have an important place in the deep learning literature [81]. This model is generally used to provide a hierarchical feature extraction for the representation of input data. It consists of a set of autoencoders, each of which is trained separately and then combined. Autoencoders perform a learned encoding operation to represent the data in a lower dimensional space and then aim to reconstruct this representation from the original data. Simple autoencoder architecture is shown in **Figure 2** [82]. Autoencoder neural network essentially consists of two elements: an encoder and a decoder. Moreover, it renovates the specified input data into features using the encoder and hence reconstructs the input data by renovating these features back to its raw data throughout the decoder [83]. Multiple autoencoders are related to each other to form a stacked autoencoder architecture. A stacked autoencoder allows learning higher-level features, each derived from the previous one. This model generally provides effective results on large and complex data sets and is widely used in feature extraction, dimensionality reduction, and classification operations [84]. It contains many hyperparameters in a stacked auto-encoder structure. These hyperparameters directly affect the performance of the established network. High success in a created architecture can be achieved by optimizing these parameters specific to the problem.
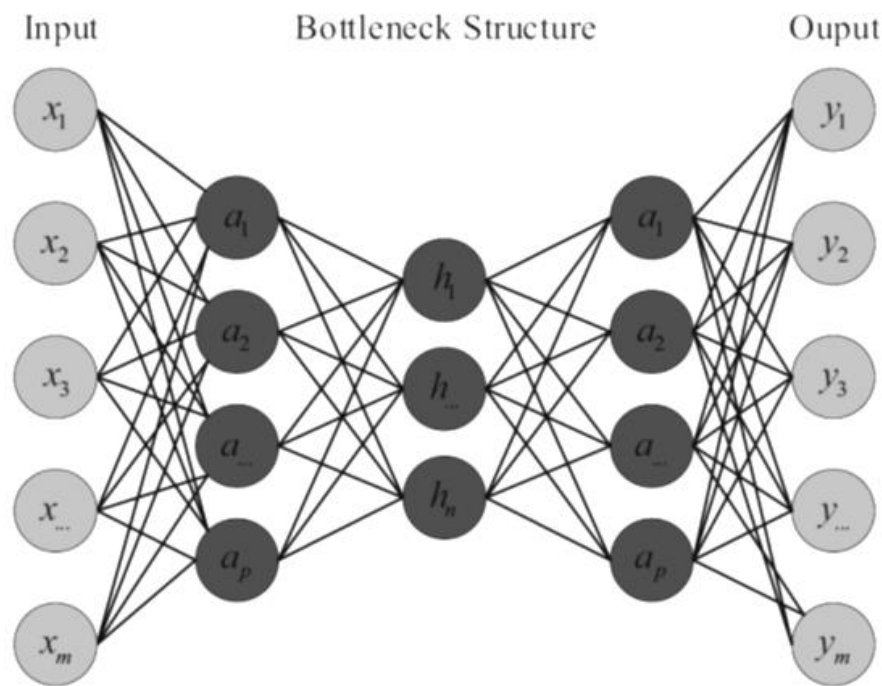

Figure 2. Auto-encoder architecture.

Stacked Autoencoders are a type of neural network used to learn hidden representations of data. Stacked autoencoders rely on stacking several layers of autoencoders to arrive at more complex representations, with each layer trained sequentially. The basic structure and mathematical equations for the stacked autoencoder network are described as follows:

(i) Encoding stage. The input x in the first layer is transformed by an activation function to generate the first hidden representation (see equation (7)).

$$h^{(1)} = f(w^1.x + b^1) \qquad (7)$$

where w^1 is the specified weight matrix related to the first layer, b^1 is the shift vector (bias), and f is the activation function such as a ReLU or Sigmoid function.

(ii) Decoding. The hidden representation $h^{(1)}$ is passed through the next layer to reconstruct the original input (see equation (8))

$$\hat{x} = g(w^2. + h^{(1)} + b^{(2)}) \qquad (8)$$

where w^2 is the specified weight matrix for the second layer, $b^{(2)}$ is the bias vector for the second layer, and g is the specified activation function used in decoding.

(iii) Cost Function. The difference between the original input x and the reconstructed input $\hat{x}$ is measured using a cost function, such as the Mean Squared Error (see equation (9)):

$$L(\text{x}, \hat{x}) = \frac{1}{n}\sum(x_i - \hat{x}_i)^2 \qquad (9)$$

where n is the number of data samples.

(iv) Stacked Layer Training. When training a stacked autoencoder, several layers are stacked so that the hidden layer is used as the input to the next layer, so the autoencoder is stacked as follows (see equation (10)):

$$h^{(2)} = f(w^3.h^1 + b^3) \qquad (10)$$

Then the encoder is decoded again to obtain the reconstructed input, and this process is repeated with each added layer.

3.3. Proposed ConvNeXt Stacked Autoencoder

A ConvNeXt Stacked Autoencoder (SAE) is an advanced neural network that integrates ConvNeXt, a modernized convolutional architecture, with a stacked autoencoder. ConvNeXt acts as the encoder, leveraging its convolutional layers to extract hierarchical, high-quality features from images, benefiting from modern techniques like large kernel sizes and layer normalization. The stacked autoencoder architecture allows for unsupervised learning, reducing dimensionality by compressing the input data into a latent space and then reconstructing it. By attaching a classifier after encoding, this architecture excels in image classification, especially where feature extraction and representation learning are critical. This combination enables improved accuracy and efficiency in image classification tasks compared to traditional CNN autoencoders, owing to ConvNeXt's improved feature extraction capabilities. The block diagram of the proposed method is shown in **Figure 3**. The steps of the ConvNeXt-SAEs algorithm are described in Algorithm 1 in **Table 2**.
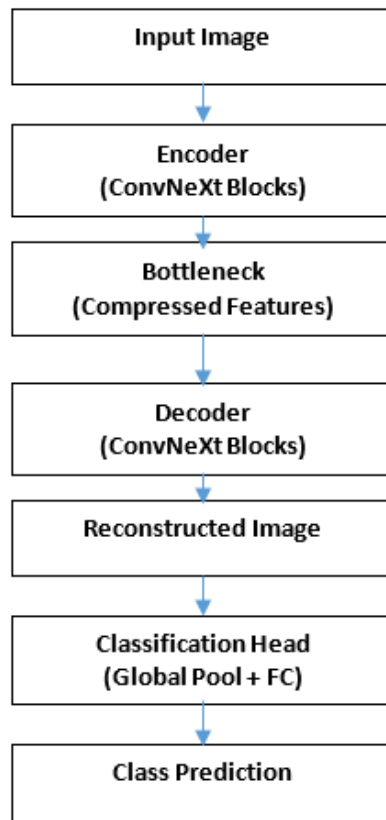
Figure 3. ConvNeXt-SAEs block diagram.

Table 2. Algorithm 1 relating to ConvNeXt-SAEs algorithm steps.

Algorithm 1. ConvNeXt-SAEs algorithm steps

1. **Data preparation:**
    *   Take a set of images {X} = {x1, x2, ..., xn}, and each image xi represents a set of pixels in the two dimensions (h×w×c), where: h: height of the image, w: width of the image and c: number of channels (e.g. 3 for RGB color channels).
2. **Create the model:**
    *   Inputs: The input image {x} is represented in the dimension (h×w×c)
3. **Encoder:**
    *   Convolution layer: The convolutional operation is applied to the image using a filter W:
    $$z_i = \boldsymbol{W_i} * \boldsymbol{x} + \boldsymbol{b_i}$$
    where z_i is output of convolution, x input image and b_i is the bias.
    *   Activation: which is a non-linear activation function is applied to the output of the specified convolutional layer, such as the ReLU function:
    $$a_i = \max(0, z_i)$$
    *   Pooling: Next, we use Max Pooling to reduce the dimensionality:
    $$p_i = \max(a_i)$$
    The output size is reduced by taking the maximum value in each window.
4. **Bottleneck:**
    *   The data is transferred to a fully connected layer. If the previous output has dimension n while the fully connected layer has m nodes, this is mathematically represented as follows:
    $$h = \sigma(W_h \, p + b_h)$$
    where *h* is the Compact representation of features, *W* Fully connected layer weight and σ is the Activation function, such as ReLU.
5. **Decoder:**
    *   Transposed Convolution Layer: To expand the dimensions again, we use UpSampling or

Table 2 (continue). Algorithm 1 relating to ConvNeXt-SAEs algorithm steps.

Algorithm 1. ConvNeXt-SAEs algorithm steps

- Transposed Convolution. The equation here is similar to the convolution layer but in reverse:

$$y_i = W_i^T * h + b_i$$

- Activation: The activation function (such as ReLU) is implemented in the same way:

$$a_i = \max(0, y_i)$$

- Upsampling: To expand the image again, the dimensions are increased using expansion operations:

$$u_i = \textbf{upsample}(a_i)$$

6. **Output Layer:**
   - A final convolutional layer to generate the reconstructed image $\hat{x}$

$$\hat{x} = \sigma(W_o * u + b_o)$$

7. **Classification Head:**
   - Global Average Pooling: The encoder output is passed to the Global Average Pooling layer which collects information from each channel:

$$g = \frac{1}{n} \sum_{i=1}^{n}(h_i)$$

   - Classification layer: The result is passed to a fully connected layer with k nodes (where k is the number of classes).

$$y = softmax(W_y g + b_y)$$

8. **Training:**
   - Use a loss function such as binary cross-entropy.

$$L(\hat{y}, y) = -\sum_{i=1}^{k} y_i \log(\hat{y}_i)$$

9. **Evaluation:**
   - Calculate the accuracy of the model based on the comparison among the predicted results y and the final actual results.

3. RESULTS AND DISCUSSION

The series of simulation results that we carried out reflect in part the major aim of the proposed ConvNeXt Stacked Autoencoders (ConvNeXt-SAEs) architecture with published data from real manufacturing lines in the Kaggle website. The industrial quality control of packages dataset contains information about package dimensions, weights, and defect indicators. It is used to train machine learning models to predict defects, detect anomalies, or improve quality control processes. This dataset is useful in industrial and academic settings to analyze and improve packaging standards through data analysis. The data includes images of packages from a real manufacturing line. The data is categorized into two types: damaged boxes, which contain 200 image files of only damaged boxes; and intact boxes, which contain 200 image files of only intact boxes. The data is split into 70% for training and 30% for testing and evaluation. Before training the proposed method preprocessing image augmentation techniques are used to generate new images from the existing ones hence increasing the used data diversity. **Figure 4** illustrates the types and the number of layers used in the proposed ConvNeXt-SAEs architecture for product defect detection. The model hyperparameters are summarized in **Table 3**. The proposed ConvNeXt-SAEs model evaluation results in terms of accuracy and model loss which are shown in **Figure 5**.

To validate the effectiveness of the proposed ConvNeXt-SAEs model for product defect detection a comparison with conventional ConvNeXt and the traditional CNN models is conducted. Both ConvNeXt and CNN methods are trained and evaluated with the same

dataset used for evaluating the ConvNeXt-SAEs model with the same parameter settings. The evaluation results of the ConvNeXt and CNN models in terms of accuracy and loss are shown in **Figures 6** and **7,** respectively. The obtained results for the proposed ConvNeXt-SAEs, ConvNeXt, and CNN models are summarized in **Table 3**.
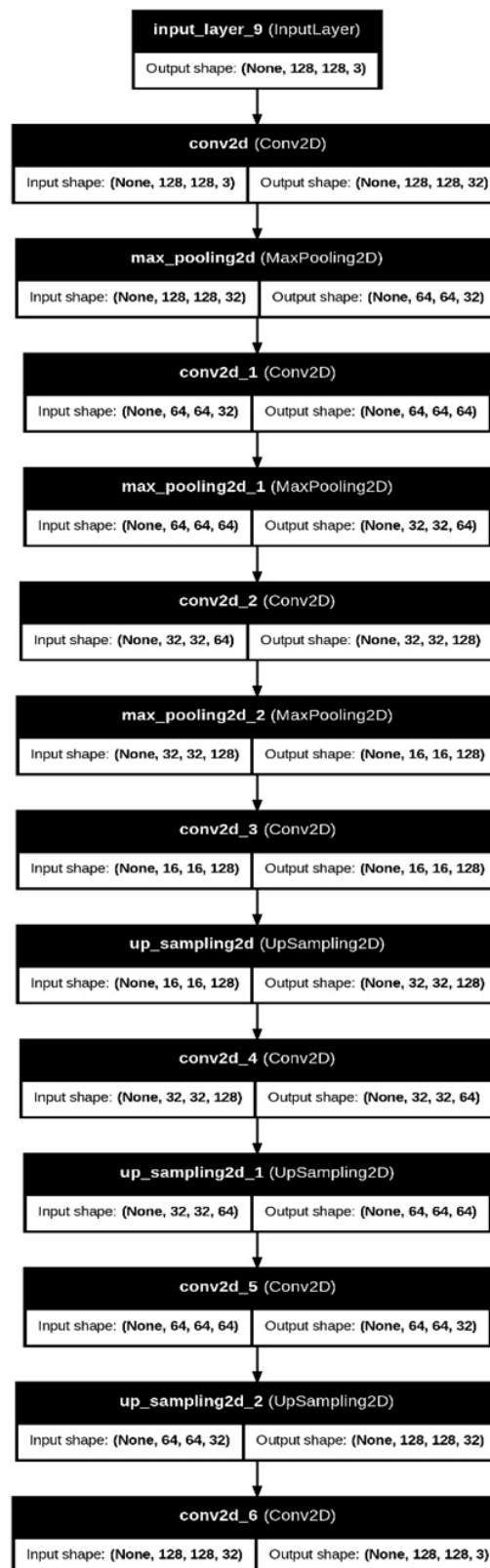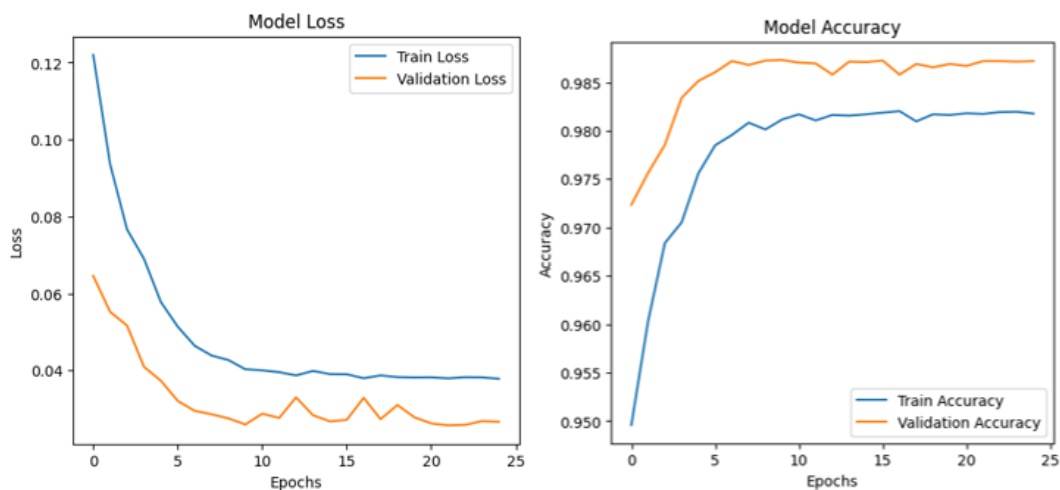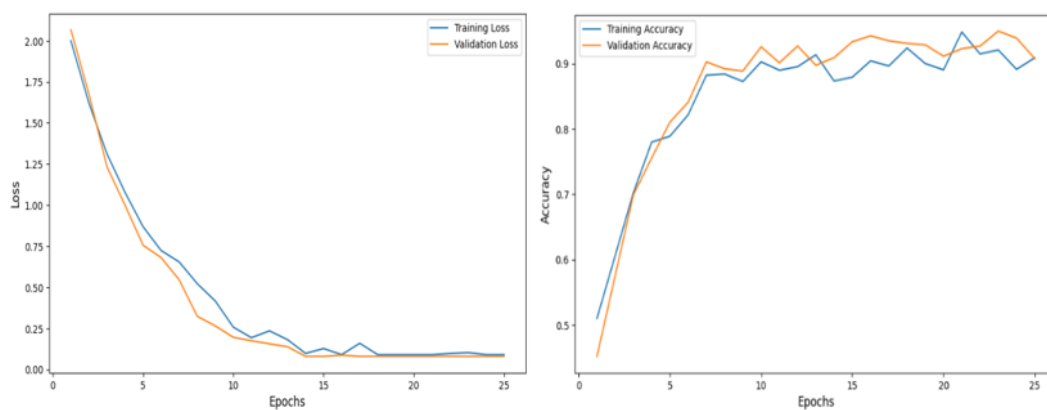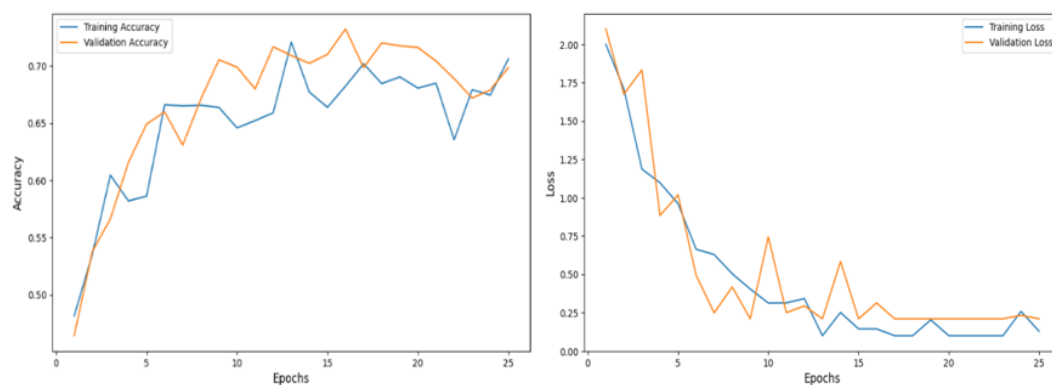

Figure 4. The proposed ConvNeXt Stacked Autoencoder architecture.

Table 3. Model hyperparameters.

Parameter	Value
Activation function	GeLU /softmax
Learning Rate	0.01
Loss function	categorical_crossentropy
Epoch	25
Batch size	8
Optimizer	ADAM

Figure 5. ConvNeXt-SAEs model results.

Figure 6. ConvNeXt model results.

Figure 7. CNN model results.

Based on the obtained results, the ConvNeXt achieves higher accuracy by 21% improvement compared to traditional convolutional CNN, thanks to architectural improvements inspired by Vision Transformers. These include using larger convolutional kernels (7x7) to capture more spatial information and replacing the Batch Normalization layer with Layer Normalization to improve model stability. It also simplifies the architecture and reduces complexity, which reduces overfitting and increases computational efficiency using deep convolutions. These improvements allow ConvNeXt to outperform CNNs on tasks such as image classification. Moreover, based on the obtained results in **Table 4**, the ConvNeXt-SAEs are more effective than ConvNeXt by 7% because of their ability to better extract complex features from data. By using multiple layers of encoder and decoder, the model can delve deeper into understanding graph representations, which helps improve accuracy in image classification or other tasks. It can also teach more complex pattern recognition to the hierarchical structure it provides. These features make ConvNeXt-SAEs a preferred choice when it comes to achieving higher accuracy in tasks that require deep data analysis.

Table 4. Model's results summary.

Method	F-Score	loss
ConvNeXt-SAEs	98.5	0.01
ConvNeXt	91.3	0.08
CNN	70.2	0.22

5. CONCLUSION

The recent advancement in the DL methods offers ways for industrial companies to meet the requirements of smart manufacturing systems due to their ability to classify and analyze the features of the input data. The use of DL algorithms in manufacturing lines has greatly decreased human involvement in production for product defect detection. The detection approach based on DL algorithms is appropriate to diverse objects and many defect types as long as it is trained well depending on the corresponding data. In this context, this paper investigates the feasibility of using DL algorithms in automated defect detection systems. The ConvNeXt Stacked Autoencoders (ConvNeXt-SAEs) DL algorithm is proposed. ConvNeXt is well-known as the most robust model that can balance between precision and recall. On the other hand, SAE can perform effective classification by analyzing the features of the input data. The proposed ConvNeXt-SAEs DL algorithm combines the core concepts of these two DL types of architecture into a single hybrid model applied to the defect-detection system. Published data from real manufacturing lines on the Kaggle website is used for evaluation. Simulation results show that the ConvNeXt-SAEs algorithm improves the accuracy by 7% and 40% as compared with the conventional ConvNeXt and CNN algorithms, respectively. Based on the obtained results, the proposed ConvNeXt-SAEs DL algorithm proves significant potential and advantages to be used in real-time industrial applications for defect detection.

6. AUTHORS' NOTE

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

7. REFERENCES

[1]   Drakaki, M., Karnavas, Y.L., Tzionas, P. and Chasiotis, I.D., (2021). Recent developments towards Industry 4.0 oriented predictive maintenance in induction motors. *Procedia Computer Science*, *180*, 943-949.

[2]   Glushchenko, V. V. (2023). Theoretical foundations of the creation of a curriculum in higher project IT in education. *Indonesian Journal of Teaching in Science*, *3*(2), 161-170.

[3]   Suyunovich, T. O., and Fayzievna, H. M. (2023). Improving training of modern leaders utilizing it in the administration of the higher education system. *Indonesian Journal of Teaching in Science*, *3*(2), 191-200.

[4]   Glushchenko, V. V. (2024). The paradigm of curriculum differentiation in higher IT education. *ASEAN Journal for Science Education*, *3*(1), 87-94.

[5]   Shaffiyah, I., Dwiyanti, V., and Masek, A. (2022). Smart city and society 5.0: Involvement of information technology in the development of public service systems in Indonesia. *ASEAN Journal of Community Service and Education*, *1*(1), 31-42.

[6]   Syahrul, S., Nurhayati, S., Wicaksono, M. F., and Fatimah, T. N. (2025). Internet of things-based child stunting detection system for supporting sustainable development goals. *Journal of Engineering Science and Technology*, *20*(2), 510-519.

[7]   Nidamanuri, S., Masood, J. A. I. S., Asirvatham, D., Abas, F. S. B., Veerappan, K. N. G., Budiman, F., and Sugumaran, V. (2025). Predictive deep learning models to identify traumatic brain injuries using MRI data. *Journal of Engineering Science and Technology*, *20*(2), 362-379.

[8]   Robandi, B., and Setiawardani, W. (2025). Enhancing digital literacy and teacher-preneurship through a critical pedagogy-based training platform. *Journal of Engineering Science and Technology*, *20*(2), 558-576.

[9]   Azhar, N. A., Radzi, N. A., Ramachandaramurthy, V. K., Abdullah, F., Kahar, N., Awang, A., and Yeep, C. K. (2025). Electric vehicle consumption dataset tailored to malaysian situation and implemented using rapid miner auto-model. *Journal of Engineering Science and Technology*, *20*(2), 490-509.

[10]   Soegoto, E. S., Rafdhi, A. A., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging applications of IoT, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis. *Journal of Engineering Science and Technology*, *20*(2), 404-426.

[11]   Anh, D. H. M. (2022). Mesh network based on MQTT broker for smart home and IIoT factory. *ASEAN Journal of Science and Engineering*, *2*(2), 173-180.

[12]   Thapwiroch, K., Kumlue, A., Saoyong, N., Taprasan, P., and Puengsungewan, S. (2021). Easy-mushroom mobile application using the Internet of Things (IoT). *Indonesian Journal of Educational Research and Technology*, *1*(1), 1-6.

[13]   Jebur, T. K. (2023). Greening the internet of things: A comprehensive review of sustainable iot solutions from an educational perspective. *Indonesian Journal of Educational Research and Technology*, *3*(3), 247-256.

[14]   Syahrul, S., Nurhayati, S., Wicaksono, M. F., and Fatimah, T. N. (2025). Internet of things-based child stunting detection system for supporting sustainable development goals. *Journal of Engineering Science and Technology*, *20*(2), 510-519.

[15]   Soegoto, E. S., Rafdhi, A. A., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging applications of IoT, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis. *Journal of Engineering Science and Technology*, *20*(2), 404-426.

[16] Luckyardi, S., Hayati, E. N., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging trends in technology: Insights across machine learning, digitalization, and industry applications. *Journal of Engineering Science and Technology*, *20*(2), 458-475.

[17] Ismail, F. B., Azahar, N. A., Dihrab, S. S., Al-Bazi, D. A., and Kazem, H. A. (2025). Internet of Things for Monitoring and Optimisation of Stand-Alone Systems in Rural Area: An Experimental Case. *Journal of Engineering Science and Technology*, *20*(2), 612-626.

[18] Huong, A. K., Sukki, J. M., and Ngu, X. T. (2025). MOBILEDFU: Classification of diabetic foot ulcer infection on the edge. *Journal of Engineering Science and Technology*, *20*(2), 577-589.

[19] Al Husaeni, D. F., Haristiani, N., Wahyudin, W., and Rasim, R. (2024). Chatbot artificial intelligence as educational tools in science and engineering education: A literature review and bibliometric mapping analysis with its advantages and disadvantages. *ASEAN Journal of Science and Engineering*, *4*(1), 93-118.

[20] Rochman, S., Rustaman, N., Ramalis, T. R., Amri, K., Zukmadini, A. Y., Ismail, I., and Putra, A. H. (2024). How bibliometric analysis using VOSviewer based on artificial intelligence data (using ResearchRabbit Data): Explore research trends in hydrology content. *ASEAN Journal of Science and Engineering*, *4*(2), 251-294.

[21] Solihat, A. N., Dahlan, D., Kusnendi, K., Susetyo, B., and Al Obaidi, A. S. M. (2024). Artificial intelligence (AI)-based learning media: Definition, bibliometric, classification, and issues for enhancing creative thinking in education. *ASEAN Journal of Science and Engineering*, *4*(3), 349-382.

[22] Rahayu, N. I., and Ismail, A. (2023). Trends in the use of artificial intelligence (ai) technology in increasing physical activity. *Indonesian Journal of Educational Research and Technology*, *3*(3), 295-304.

[23] Fiandini, M., Nandiyanto, A. B. D., and Kurniawan, T. (2023). Bibliometric analysis of research trends in conceptual understanding and sustainability awareness through artificial intelligence (AI) and digital learning media. *Indonesian Journal of Multidiciplinary Research*, *3*(2), 477-486.

[24] Nurhasanah, S., and Nugraha, M. S. The future of learning: Ethical and philosophical implications of artificial intelligence (AI) integration in education. *Indonesian Journal of Multidiciplinary Research*, *4*(2), 341-352.

[25] Alimi, A. E., Buraimoh, O. F., Aladesusi, G. A., and Babalola, E. O. (2021). University students' awareness of, access to, and use of artificial intelligence for learning in Kwara State*, Indonesian Journal of Teaching in Science, 1*(2), 91-104.

[26] Rahmiyanti, H. (2024). Bibliometric analysis on artificial intelligence research in Indonesia vocational education. *ASEAN Journal for Science Education*, *3*(2), 183-192.

[27] Agarry, R. O., Omolafe, E. V., Animashaun, V. O., and Babalola, E. O. (2022). Primary education undergraduates' competency in the use of artificial intelligence for learning in Kwara State. *ASEAN Journal of Educational Research and Technology*, *1*(2), 111-118.

[28] Muhabbat, H., Mukhiddin, K., Jalil, H., Dustnazar, K., Farxod, T., Shavkat, M., Khulkar, K., and Jakhongir, S. (2024). The digital frontier: AI-enabled transformations in higher education management. *Indonesian Journal of Educational Research and Technology*, *4*(1), 71-88.

[29] Viorennita, A., Dewi, L., and Riyana, C. (2023). The role of ChatGPT AI in student learning experience. *Indonesian Journal of Multidiciplinary Research*, *3*(2), 445-452.

[30] Abulude, F. O., Acha, S., Arifalo, K. M., Akinnusotu, A., Oluwagbayide, S. D., Al Shoumik, B. A., and Gbotoso, A. O. (2023). Monitoring of air quality with satellite-based sensor:

The case of four towns in Southeast, Nigeria. *ASEAN Journal of Science and Engineering*, *3*(1), 39-46.

[31] Ratsame, P., Koedsaksit, M., Jitsoontron, S., Jaisamak, U., and Puengsungewan, S. (2021). Vehicle-tracking mobile application without a GPS sensor. *Indonesian Journal of Educational Research and Technology*, *1*(1), 11-14.

[32] Jebur, T. K. (2024). Securing wireless sensor networks, types of attacks, and detection/prevention techniques, an educational perspective. *ASEAN Journal of Science and Engineering Education*, *4*(1), 43-50.

[33] Dilshad, M., Hussain, T., Ali, S., Shahzad, M. S., Javed, M., Ahmad, F. N., and Zaki, A. A. (2025). Graphene-based electrochemical sensors for heavy metal ions detection: A comprehensive review. *ASEAN Journal for Science and Engineering in Materials*, *4*(1), 23-42.

[34] Syahrul, S., Nurhayati, S., Wicaksono, M. F., and Fatimah, T. N. (2025). Internet of things-based child stunting detection system for supporting sustainable development goals. *Journal of Engineering Science and Technology*, *20*(2), 510-519.

[35] Luckyardi, S., Hayati, E. N., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging trends in technology: Insights across machine learning, digitalization, and industry applications. *Journal of Engineering Science and Technology*, *20*(2), 458-475.

[36] Taha, I. Z., and Al Janaby, A. O. (2025). Interference mitigation for dynamic user connectivity using sdn and radio resource management in cell-less networks. *Journal of Engineering Science and Technology*, *20*(2), 520-535.

[37] Ismail, F. B., Azahar, N. A., Dihrab, S. S., Al-Bazi, D. A., and Kazem, H. A. (2025). Internet of things for monitoring and optimisation of stand-alone systems in rural area: An experimental case. *Journal of Engineering Science and Technology*, *20*(2), 612-626.

[38] Huong, A. K., Sukki, J. M., and Ngu, X. T. (2025). MOBILEDFU: Classification of diabetic foot ulcer infection on the edge. *Journal of Engineering Science and Technology*, *20*(2), 577-589.

[39] Bangkerd, P., and Sangsawang, T. (2021). Development of augmented reality application for exercise to promote health among elderly. *Indonesian Journal of Educational Research and Technology*, *1*(3), 77-80.

[40] Albar, C. N., Widiansyah, M. G., Mubarok, S., Aziz, M. A., and Maulana, H. (2021). Application of augmented reality technology with the fuzzy logic method as an online physical education lecture method in the new normal era. *Indonesian Journal of Multidiciplinary Research*, *1*(1), 35-40.

[41] Al Husaeni, D. N., Munir, M., and Rasim, R. How to create Augmented Reality (AR) applications using unity and Vuforia engine to teach basic algorithm concepts: Step-by-step procedure and bibliometric analysis. *Indonesian Journal of Teaching in Science*, *4*(2), 189-204.

[42] Soegoto, E. S., Rafdhi, A. A., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging applications of iot, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: a comprehensive review and analysis. *Journal of Engineering Science and Technology*, *20*(2), 404-426.

[43] Luckyardi, S., Hayati, E. N., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging trends in technology: insights across machine learning, digitalization, and industry applications. *Journal of Engineering Science and Technology*, *20*(2), 458-475.

[44] Almurtadha, Y. (2025). Let's immerse in metaverse: Overview, challenges, and stumeta framework for successful implementation. *Journal of Engineering Science and Technology*, *20*(2), 380-403.

[45] Istianah, F., Hidayat, A., Handayanto, S. K., Utama, C., Susilawati, A., and Abadi, S. (2025). Development of science virtual laboratory (scivlab) to develop critical thinking skills in elementary schools on the topic of changes in the state of substances. *Journal of Engineering Science and Technology, 20*(1), 9-16.

[46] Meirawan, D., Muktiarni, M., Rahayu, S., Sabitri, Z., and Ghinaya, Z. (2025). Usability assessment of flipo-ar: navigating learning in a vuca world with augmented reality. *Journal of Engineering Science and Technology, 20*(1), 137-144.

[47] Firdiarahma, F. (2021). The use of virtual reality as a substitute for the pre-school students' field trip activity during the learning from home period. *Indonesian Journal of Educational Research and Technology*, *1*(2), 57-60.

[48] Ekunola, G. T., Onojah, A. O., Talatu, A. F., and Bankole, M. O. (2022). Colleges of education lecturers' attitude towards the use of virtual classrooms for instruction. *Indonesian Journal of Multidiciplinary Research*, *2*(1), 187-194.

[49] Bugarso, J. M. S., Cabantugan, R. E., Que-ann, D. T., and Malaco, A. C. (2021). Students' learning experiences and preference in performing science experiments using hands-on and virtual laboratory. *Indonesian Journal of Teaching in Science*, *1*(2), 147-152.

[50] Azizah, E. V., Nandiyanto, A. B. D., Kurniawan, T., and Bilad, M. R. (2022). The effectiveness of using a virtual laboratory in distance learning on the measurement materials of the natural sciences of physics for junior high school students. *ASEAN Journal of Science and Engineering Education*, *2*(3), 207-214.

[51] Salman, A. A., and Yahaya, O. (2025). Perception of early childhood education lecturers on the use of virtual learning. *ASEAN Journal for Science Education*, *4*(1), 31-38.

[52] Ekunola, G. T., Obielodan, O. O., and Babalola, E. O. (2022). Lecturers perceived proficiency in the use of virtual classrooms for instruction in colleges of education. *ASEAN Journal of Educational Research and Technology*, *1*(1), 7-16.

[53] Sison, A. J. R. N., Bautista, J. M., Javier, J. L., Delmonte, R. J. B., and Cudera, R. B. (2024). Development and acceptability of virtual laboratory in learning systematics. *ASEAN Journal of Educational Research and Technology*, *3*(1), 9-26.

[54] Rivky, M., Fajar, M. R. K., and Pangestu, A. R. (2022). Utilization of virtual reality chat as a means of learning communication in the field of education. *ASEAN Journal of Community Service and Education*, *1*(1), 23-30.

[55] Azizah, E. V., Nandiyanto, A. B. D., Kurniawan, T., and Bilad, M. R. (2022). The effectiveness of using a virtual laboratory in distance learning on the measurement materials of the natural sciences of physics for junior high school students. *ASEAN Journal of Science and Engineering Education*, *2*(3), 207-214.

[56] Chen, X. R., Ku, H. H., Chen, Y., Chen, W., Dai, S., Lin, C. J., Yu, M. H., and Weng, T. S. (2025). Developing an identification and testing system for cardiac surgical instrument capabilities. *Journal of Engineering Science and Technology*, *20*(2), 476-489.

[57] Soegoto, E. S., Rafdhi, A. A., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging applications of IOT, machine learning, virtual reality, augmented reality and artificial intelligent in monitoring systems: A comprehensive review and analysis. *Journal of Engineering Science and Technology*, *20*(2), 404-426.

[58] Luckyardi, S., Hayati, E. N., Abduh, A., Rosmaladewi, R., and Haristiani, N. (2025). Emerging trends in technology: Insights across machine learning, digitalization, and industry applications. *Journal of Engineering Science and Technology*, *20*(2), 458-475.

[59] AlmurtadhA, Y. (2025). Let's immerse in metaverse: overview, challenges, and stumeta framework for successful implementation. *Journal of Engineering Science and Technology*, *20*(2), 380-403.

[60] Istianah, F., Hidayat, A., Handayanto, S. K., Utama, C., Susilawati, A., and Abadi, S. (2025). Development of science virtual laboratory (SCIVLAB) to develop critical thinking skills in elementary schools on the topic of changes in the state of substances. *Journal of Engineering Science and Technology, 20*(1), 9-16.

[61] Wang, T., Chen, Y., Qiao, M. and Snoussi, H., (2018). A fast and robust convolutional neural network-based defect detection model in product quality control. *The International Journal of Advanced Manufacturing Technology*, *94*, 3465-3471.

[62] Yang, J., Li, S., Wang, Z., Dong, H., Wang, J. and Tang, S., (2020). Using deep learning to detect defects in manufacturing: A comprehensive survey and current challenges. *Materials*, *13*(24), 5755.

[63] Villalba-Diez, J., Schmidt, D., Gevers, R., Ordieres-Meré, J., Buchwitz, M. and Wellbrock, W., (2019). Deep learning for industrial computer vision quality control in the printing industry 4.0. *Sensors*, *19*(18), 3987.

[64] Abood, A.M., Nasser, A.R. and Al-Khazraji, H., (2023). Predictive maintenance of electromechanical systems based on enhanced generative adversarial neural network with convolutional neural network. *International Journal of Artificial Intelligence Research, 2252*, 8938.

[65] Al-Khazraji, H., Nasser, A.R. and Khlil, S., (2022). An intelligent demand forecasting model using a hybrid of metaheuristic optimization and deep learning algorithm for predicting concrete block production. *IAES International Journal of Artificial Intelligence*, *11*(2), 649.

[66] Nasser, A. and Sever, H., (2020). A concept-based sentiment analysis approach for Arabic. *The International Arab Journal of Information Technology*, *17*(5), 778-788.

[67] Cardoso, D. and Ferreira, L., (2020). Application of predictive maintenance concepts using artificial intelligence tools. *Applied Sciences*, *11*(1), 18.

[68] Nasser, A. and Al-Khazraji, H., (2022). A hybrid of convolutional neural network and long short-term memory network approach to predictive maintenance. *International Journal of Electrical and Computer Engineering (IJECE)*, *12*(1), 721-730.

[69] Abood, A.M., Nasser, A.R. and Al-Khazraji, H., (2022). Predictive maintenance of electromechanical systems using deep learning algorithms. *Ingénierie des Systèmes d'Information*, *27*(6), p1009.

[70] Ozdemir, R. and Koc, M., (2019), September. A quality control application on a smart factory prototype using deep learning methods. *In 2019 IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT), 1*, 46-49.

[71] Song, S., Jing, J., Huang, Y. and Shi, M., (2021). EfficientDet for fabric defect detection based on edge computing. *Journal of Engineered Fibers and Fabrics*, *16*, 15589250211008346.

[72] Chaabi, M., Hamlich, M. and Garouani, M., (2023). Product defect detection based on convolutional autoencoder and one-class classification. *IAES International Journal of Artificial Intelligence*, *12*(2), 912.

[73] Zheng, X., Zheng, S., Kong, Y. and Chen, J., (2021). Recent advances in surface defect inspection of industrial products using deep learning techniques. *The International Journal of Advanced Manufacturing Technology*, *113*, 35-58.

[74] Rahman, N.A., Saad, N.M., Abdullah, A.R., Hassan, M.R.M., Basir, M.S.S.M. and Noor, N.S.M., (2018). Automated real-time vision quality inspection monitoring system. *Indonesian Journal of Electrical Engineering and Computer Science*, *11*(2), 775-783.

[75] Tercan, H. and Meisen, T., (2022). Machine learning and deep learning based predictive quality in manufacturing: A systematic review. *Journal of Intelligent Manufacturing*, *33*(7), 1879-1905

[76] Likhosherstov, V., Choromanski, K. and Weller, A., (2023), June. On the expressive flexibility of self-attention matrices. *In Proceedings of the AAAI Conference on Artificial Intelligence, 37*(7), 8773-8781.

[77] Li, Z., Gu, T., Li, B., Xu, W., He, X. and Hui, X., (2022). ConvNeXt-based fine-grained image classification and bilinear attention mechanism model. *Applied Sciences*, *12*(18), 9016.

[78] Aksoy, S., Demircioglu, P. and Bogrekci, I., (2024). Enhancing Melanoma Diagnosis with Advanced Deep Learning Models Focusing on Vision Transformer, Swin Transformer, and ConvNeXt. *Dermatopathology*, *11*(3), 239-252.

[79] Liu, Z., Mao, H., Wu, C.Y., Feichtenhofer, C., Darrell, T. and Xie, S., (2022). A convnet for the 2020s. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022,* 11976-11986.

[80] Reyes, D. and Sánchez, J., (2024). Performance of convolutional neural networks for the classification of brain tumors using magnetic resonance imaging. *Heliyon*, *10*(3), e25468.

[81] Kim, J.Y. and Oh, J.S., (2023). A Pilot study of stacked autoencoders for ship mode classification. *Applied Sciences*, *13*(9), 5491.

[82] Bao, W., Yue, J., and Rao, Y., (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, *12*(7), e0180944.

[83] Schuster, V. and Krogh, A., (2021). A manifold learning perspective on representation learning: learning decoder and representations without an encoder. *Entropy*, *23*(11), 1403.

[84] Hu, X., Yin, Z., Zeng, Z. and Peng, Y., (2023). Prediction of miRNA–Disease associations by cascade forest model based on stacked autoencoder. *Molecules*, *28*(13), 5013.