Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 16 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i16.53159

PAPER

Integrating Human-Computer Interaction and Software Engineering for Enhanced Usability Using Support Vector Machines

Jasem Alostad(✉)

The Public Authority of Applied Education and Training, Kuwait, Kuwait

jm.alostad@paaet.edu.kw

ABSTRACT

Software system development is evolving as human-computer interaction (HCI) concepts find application in software engineering. This shift emphasises more the need of knowing user requirements, behaviours, and preferences. This approach is especially important to meet the always rising demand for straightforward-to-run and-understand software systems. On the other hand, traditional methods of software development sometimes ignore the need for usability, which generates products that satisfy less than user expectations. This work suggests a new framework combining HCI methods with software engineering techniques based on the results. Support vector machine (SVM) application helps to improve process usability in development. There are many iterative prototyping, usability testing, and constant user feedback components inside the framework proposed for the software development lifecycle. Strong machine learning (ML) techniques, SVM, are applied for modelling and prediction of user satisfaction based on interaction data acquired during usability testing. Data analysis helps the SVM model find significant factors affecting usability. This enables developers to make sensible decisions exactly matched with consumer expectations. The results of the pertinent tests revealed the degree of improvement in software usability this approach brought about. Apart from a precision of 90.8%, a recall of 91.7%, and an F1-score of 91.2%, the SVM model was evaluated to have reached a classification accuracy of 92.3%. HCI, combined with software engineering, achieved by means of SVM, offers a strong approach for the development of software that not only serves but also strongly connects with users, resulting in a more significant impact on society.

KEYWORDS

human-computer interaction (HCI), software engineering, usability, support vector machines (SVMs), user-centred design

Alostad, J. (2025). Integrating Human-Computer Interaction and Software Engineering for Enhanced Usability Using Support Vector Machines. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(16), pp. 41–59. https://doi.org/10.3991/ijim.v19i16.53159

Article submitted 2024-11-02. Revision uploaded 2025-04-09. Final acceptance 2025-06-18.

1 INTRODUCTION

Artificial intelligence (AI) integration into the healthcare industry, particularly in the field of gastrointestinal (GI) endoscopy, is revolutionising both diagnosis and treatment approaches. Two AI technologies, machine learning (ML) and convolutional neural networks (CNN) [1, 2] are progressively being applied to improve endoscopic operations from patient preparation to the detection of GI pathologies in real time. Given the complexity of endoscopic operations and the enormous volume of data generated during these interventions, healthcare professionals need access to robust AI solutions that can enable them in making accurate and timely decisions [3].

Although AI has great potential, several elements prevent this general acceptance of it in GI endoscopy. It is important to underline that differences between human cognition and AI's decision-making capacity might lead to problems like automation bias, in which doctors might rely too much on AI recommendations, so neglecting their own knowledge [4]. AI systems generating too many alarms can cause alarm fatigue, a condition that can desensitise medical professionals and cause their possible neglect of important results [5]. The phenomenon of algorithm aversion, a resistance to trusting the results of AI, adds to the complications of the effective use of these technologies. These kinds of challenges influence not only the general efficiency of AI systems but also raise issues about patient safety and the usability of AI in medical environments [7].

This work attempts to address the optimization of human-AI interaction (HAII) inside the framework of GI endoscopy by means of mathematical analysis. It is still mostly unknown how best to include these technologies into clinical procedures without compromising diagnosis accuracy or clinician confidence [8]. Although the use of AI in this domain has shown that it has good potential. Often producing high false positive rates, the methods under use can result in non-needed treatments and increase healthcare costs [9]. Lack of user-centred design is another factor affecting usability problems in AI interfaces; this so lowers suitable clinician involvement and acceptance [10], [11]. As is quite evident, there is a great need for an AI system that strikes a balance between technological capacity and human values. This highlights the need of developing solutions improving diagnosis procedures concerning usability and accuracy [12].

The main objective of this work is to develop and evaluate AI algorithms based on DenseNet to minimise false positives in GI endoscopy and concurrently maximise the experience of interacting with AI. The study intends to improve the usability of AI interfaces by means of user-centred design ideas, thus fostering a more effective cooperation between doctors and such systems. Unlike other research, this one stresses the efficiency of technology combined with the factors affecting human behaviour.

The novelty of this study lies in tackling the difficult problems concerning HAII in GI endoscopy from a new angle. Apart from using modern DenseNet architectures to improve diagnosis accuracy, the proposed approach aims to effectively involve doctors by means of easily comprehensible AI interfaces. By means of empirical data on the positive effects of user-centred design on clinician performance outcomes and satisfaction, this work advances the field. Moreover, the studies provide a framework to evaluate the interaction between AI performance criteria and clinician experiences. This opens the path for the next AI-driven medical technology developments.

2 RELATED WORKS

Application of AI in GI endoscopy has developed rapidly, allowing the resolution of significant challenges and concurrently generating new scientific ideas. Using video capsule endoscopy (VCE), Wahab et al. [11] investigated the potential discrepancies between the clinically relevant recommendations created by gastroenterologists and the ML research methods now under use. By means of their work, they proposed a taxonomic hierarchy and contextually relevant interpretation of raw VCE frames into results. Their goal was to use a structured mapping of frame-level results to disease-level diagnosis, thereby improving classification metrics and enhancing discriminative learning.

Campion et al. [12] became interested in the expanding range of AI applications in GI endoscopy. They also underlined the need for having consistent ML and CNN platforms to support clinical interventions. Their argument is that interactions between doctors and AI technologies are shaped by psychological and behavioural aspects, so generating complexity in HAII. Reducing false positives and building easily used interfaces are two of the most important components of optimising HAII, the authors observed. By stressing significant human factors influencing HAII, including automation bias, alarm fatigue, and algorithm aversion, they underlined the need for a human-centred design approach in the evolution of new AI technologies.

Parallel in nature, Parasa et al. [13] provided a framework for endoscopy AI ecosystem players. The framework concentrated on the criteria, measurements, and evaluation strategies essential for the clinical acceptance and implementation of AI applications. Their efforts serve as a compass, compiling ideal criteria for evaluating developing AI technologies. They emphasize the need of constant updates to show field developments during this process.

For their viability in clinical settings, Shahab et al. [14] examined large language models (LLMs) meant especially for gastroenterologists. Within the field of clinical gastroenterology, they examined possible uses, underlined their main benefits and disadvantages, and examined the technical foundations of LLMs. Their study sought to close the distance between AI technology and clinical practice by providing gastroenterologists with the required knowledge to enable active participation in the development and implementation of LLMs. Their analysis also addressed significant challenges in implementation.

Gong et al. [15] developed a clinical decision support system (CDSS) based on deep learning. Under real-time endoscopy, this system seeks to automatically recognise and categorise gastric neoplasms. The researchers obtained a lesion detection rate of 95.6% in internal tests by means of AI-assisted endoscopy, so demonstrating the advantage of this technique over more conventional ones. With a classification accuracy of 81.5%, the CDSS indicates that it could improve diagnosis capacity in clinical settings for a spectrum of gastric neoplasms.

Rodrigues Barbosa et al. [16] made a significant contribution to the field by applying a deep learning algorithm able of classifying GI organs to capsule endoscopy (CE) videos on a platform free of coding. Their model let them achieve incredible performance measures: an overall accuracy of 0.98, precision of 0.89, recall of 0.97, and F1 score of 0.92. They verified their model against several lesions using a big dataset of CE videos, so demonstrating the potential value of the model in clinical practice. Apart from improving the interpretability of the prediction of the model,

the new method of visualising the transitional area of every GI organ offers a simpler knowledge of the endoscopic results.

Based on pFMG, [17] demonstrate how flexible and capable the human-machine interface (HMI) armband is to interact with and control robotic prosthetic hands, drones, computer games, and any other system in which humans are involved. By means of an ML algorithm, which generates control signals, one can decode and classify the acquired bio signals of various hand gestures, so enabling fast and accurate recognition of the intentions of a user. [18] A multi-information recognition human-computer interaction (HCI) system has been developed using the outstanding PCPS and TENG performance. This system detects the pressure and surface material of the balls simultaneously and allows real-time interaction with a virtual character. Presenting a new double-network conductive hydrogel in [19], single-pot synthesis of lignosulfonate sodium and ionic liquid was achieved. The capacity of the gel to withstand mechanical strain for HCI. Apart from good biocompatibility, [20] showing exceptional heal ability due to its high healing efficiency, strong recyclability, and consistent antibacterial capacity.

These studies, which underline a spectrum of approaches combining ML, user-centred design, and real-time data processing [21], so reflect the rapid advances in AI technologies for GI endoscopy. Though the research under progress shows a deliberate attempt to create solutions that raise diagnosis accuracy and increase clinician interaction with AI technologies [22–24], still major challenges are HAII, data shortage, and algorithmic bias. Future research should give advanced algorithms together with strong human factors considerations top priority in order to ensure that AI technologies effectively complement clinical processes and improve patient outcomes. This is so true since the field is still under development.

Table 1. Summary of related works

Method	Algorithm	Methodology	Outcomes
Wahab et al. [11]	ML-based Framework	Interpreting raw VCE frames into findings using a taxonomical hierarchy	Improved classification metrics; need for large datasets
Campion et al. [12]	CNNs	Evaluating AI applications and HAII factors in GI endoscopy	Highlighted complexities in HAII; emphasized user design
Parasa et al. [13]	AI Standards Framework	Guidelines for metrics and evaluation methods for AI applications in endoscopy	Enhanced adoption of AI technologies
Shahab et al. [14]	LLMs	Overview of technical foundations and clinical applications of LLMs in gastroenterology	Identified barriers to implementation
Gong et al. [15]	Deep Learning (CDSS)	Automated detection and classification of gastric neoplasms during endoscopy	95.6% lesion detection rate; improved diagnostic accuracy
Rodrigues Barbosa et al. [16]	Deep Learning (Classification)	Classifying GI organs from CE videos using a no-code platform	High accuracy metrics; visualization of transitional areas

While AI applications for GI endoscopy have made tremendous progress, much study is still needed to maximise HAII to increase usability while concurrently reducing the false positive count. Currently available research mostly concentrates on algorithm performance and diagnostic accuracy and sometimes ignores the psychological and behavioural factors influencing clinician interaction with AI tools. Moreover, various models combine user-centred design ideas with sophisticated AI technologies. From this, better patient outcomes and higher degrees of clinician satisfaction could follow.

3 PROPOSED HUMAN-AI INTERACTION IN GI ENDOSCOPY

The proposed method is designed for real-time analysis of GI endoscopic images that generates and executes an algorithm based on DenseNet specialisation. The primary objective is to simultaneously solve HAII problems by means of an AI platform including user-centred design concepts, greatly improving diagnosis accuracy as illustrated in Figure 1.

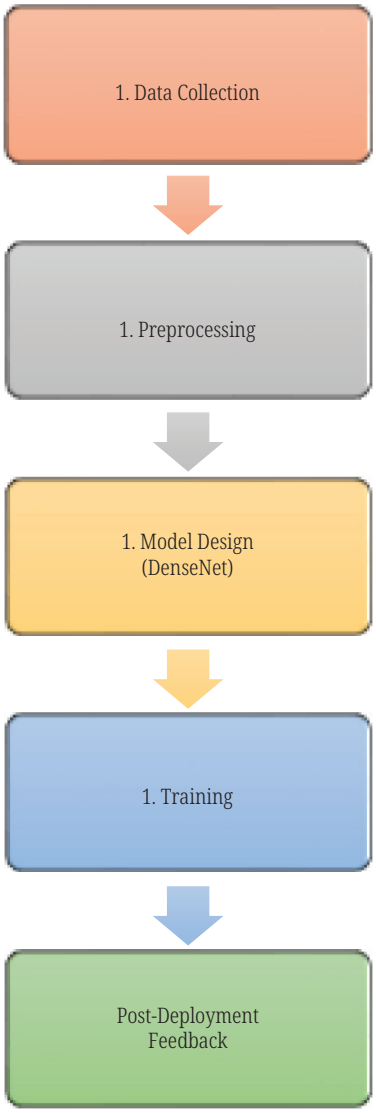

Fig. 1. Proposed optimizing human-AI interaction in GI endoscopy

Pseudocode:

```
# Initialize DenseNet model with pretrained weights
model = DenseNet(pretrained=True)
# Preprocessing step
def preprocess_image(image):
    # Apply normalization
    image = normalize(image)
    # Perform data augmentation for variability
    augmented_image = augment(image)
    return augmented_image
# Training process
def train_model(model, training_data, labels):
    optimizer = Adam() # Optimizer for minimizing loss
    loss_function = CrossEntropyLoss()
    for epoch in range(max_epochs):
        for image, label in training_data:
            preprocessed_image = preprocess_image(image)
            prediction = model(preprocessed_image)
            loss = loss_function(prediction, label)
            optimizer.zero_grad()
            loss.backward() # Backpropagation
            optimizer.step() # Update weights
# Evaluation function
def evaluate_model(model, test_data):
    false_positive_rate = 0
    for image, label in test_data:
        prediction = model(image)
        if is_false_positive(prediction, label):
            false_positive_rate += 1
    return false_positive_rate / len(test_data)
# Implement user-centered design elements
def user_centered_interface():
    # Simplicity and usability enhancements
    create_dashboard(minimalist_design=True)
    enable_contextual_feedback(for_alarm_fatigue=True)
# Final deployment
trained_model = train_model(model, training_data, labels)
fpr = evaluate_model(trained_model, test_data)
user_interface = user_centered_interface()
```

3.1 Data collection

The data collection phase helps to gather and organise a good number of GI endoscopic images. This dataset will be the basis of training and evaluation of the DenseNet-based AI model. This phase seeks to compile varied and high-quality data so that the AI system may extend over a broad spectrum of patient situations and conditions. This knowledge will cover several GI disorders, including polyps, ulcers, and inflammation.

An image acquisition. The data sources used in these contexts come from clinical settings including hospitals and GI clinics where endoscopic procedures are carried out. Images are obtained using endoscopes since they can create highly resolved videos and images of the GI tract. The dataset Denoted by the notation $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, consists of N labelled examples, Where:

- $x_i \in \mathbb{R}^{H \times W \times C}$ represents an endoscopic image, with height H, width W, and C colour channels (typically RGB).

- $y_i \in \{0, 1\}$ is the label corresponding to the image, indicating whether the image contains a pathological finding (e.g., 1 for a polyp, 0 for normal tissue).

Multi-class labels could be included in the dataset to divide among several types of pathologies, including ulcers, polyps, and inflammation. The label y_i can thus be extended to $y_i \in \{0, 1, 2, ..., K–1\}$ for a dataset including K classes.

Data annotation. Gastroenterologists, subject-matter experts, have annotated every image in the dataset, ensuring accurate supervision during the DenseNet model's training and providing ground truth labels. Assume $\boldsymbol{Y}$ represents the ground truth labels for the dataset; hence, the following is reasonable:

$$\boldsymbol{Y} = [y_1, y_2, ..., y_N]$$

Where:
y_i – element or label corresponding to the image x_i.

Data augmentation and balancing. Given some pathological cases are rarer than others, the dataset may suffer from class imbalance, so biased the model could be during training. This is supposed to be resolved by data augmentation, or the artificial increase of under-represented class size. $T_k(x_i)$ transforms every image to produce extra samples. Among several augmenting techniques, the notation k calls for zooming, flipping, and rotation.

$$x_i' = T_k(x_i)$$

The expanded dataset, $\mathcal{D}'$, then becomes:

$$\mathcal{D}' = \left\{(T_k(x_i), y_i) \mid x_i \in \mathcal{D}, k \in \left\{1,2,\ldots,M\right\}\right\}$$

Where:
M – number of augmentations applied.
This helps to balance the dataset among all the classes, so reducing the bias in the predictions of the model.

Pre-processing and normalization. Before the DenseNet model feeds the images, pre-processing methods are applied to them to standardise the data. Usually shown by the $H' \times W'$, every image is resized as part of this process to a predefined resolution. Additionally, normalised are the pixel values to show a range of [0, 1]. Using the following computation, one may derive the normalised image x_i^{norm}:

$$x_i^{\text{norm}} = \frac{x_i - \mu}{\sigma}$$

Where:
μ – mean of pixel intensity across the dataset and
σ – standard deviation.
This normalising stage guarantees consistency of the DenseNet model's inputs, promoting improved convergence during the training process.

Splitting into training, validation, and test sets. The dataset is split into three subsets after complete preparation: the training set $\mathcal{D}_{\text{train}}$, the validation set $\mathcal{D}_{\text{val}}$, and the test set $\mathcal{D}_{\text{test}}$. Usually, stratified sampling helps one to preserve the class distribution over all sets by means of these sets. The following notation allows one to demonstrate the extent of every subset: $|\mathcal{D}_{\text{train}}|$ for the number of training samples, $|\mathcal{D}_{\text{val}}|$

for the number of validation samples, and $|\mathcal{D}_{\text{test}}|$ for the number of test samples. The split is expressed as:

$$\mathcal{D} = \mathcal{D}_{\text{train}} \cup \mathcal{D}_{\text{val}} \cup \mathcal{D}_{\text{test}}$$

such that $\mathcal{D}_{\text{train}} \cap \mathcal{D}_{\text{val}} \cap \mathcal{D}_{\text{test}} = \varnothing$

With this methodical approach to data collecting, one guarantees that the DenseNet model is trained on a balanced, varied, high-quality dataset. This optimises GI endoscopy, the generalisability of the model, as well as its accuracy.

3.2 Pre-processing

Pre-processing is a crucial step in getting the obtained GI endoscopic images ready for DenseNet-based AI training. This phase aims to standardise the data entering the system, improve the image quality, and guarantee that they fit the CNN to learn from effectively. Usually including data augmentation, normalisation, resizing and normalisation, where preprocessing consists of several very important steps.

Image resizing. Before beginning the pre-processing process, all of the endoscopic images must be resized to a fixed resolution; this is usually stated as $H' \times W'$. Such homogeneity is necessary since neural networks depend on inputs with compatible dimensions. There is one practical mathematical definition for the resizing process as follows:

$$x_i^r = R\left(x_i, H', W'\right)$$

where x_i – original image.

This lets x_i pixels be resampled such that their dimensions match those given. Resizing guarantees that the basic features of the image are kept and helps to preserve a constant aspect ratio, so avoiding distortion.

Normalization. The pixel values of the images must be normalised both to enable faster convergence during training and to reduce the effect of various lighting conditions once they have been resized. Usually starting from their original range, typically [0,255], to a range of [0, 1] for 8-bit images. In mathematical terms, this is:

$$x_i^{\text{norm}} = \frac{x_i - \mu}{\sigma}$$

By normalising the images, one can help to solve issues with different lighting and contrast settings. This enables the DenseNet model to concentrate on structural elements of the images instead of illumination-induced variations.

Data augmentation. Data augmentation techniques are used to raise the resilience of the model and reduce the overfitting development possibility. Common augmenting techniques are rotation, translation, flipping, and scaling below. One can show every change depending on the type of augmentation being done using a function called T_k:

$$x_i^{\text{aug}} = T_k(x_i^{\text{norm}})$$

where K denotes the type of augmentation

The augmented dataset is formulated as:

$$\mathcal{D}^{\text{aug}} = \left\{(T_k(x_i^{\text{norm}}), y_i) \mid x_i \in \mathcal{D}\right\}$$

This change brings variability as well as increases the effective size of the dataset. This variability increases the model's generalising capacity by helping it to learn to identify traits over a range of orientations and conditions.

Image standardization. Apart from normalising images, standardising them helps to ensure that the pixel values are centred on zero and have a variance of like one unit. One could project this mathematically as follows:

$$X_i^{\text{std}} = \frac{X_i^{\text{norm}} - \mu_{\text{train}}}{\sigma_{\text{train}}}$$

Where:
μ_{train} – computed mean from the training dataset and
σ_{train} – standard deviation from the training dataset.

Standardising has the possible to improve the performance of the model even more by ensuring that the gradient descent optimization process maintains acting in a more predictable manner during training.

Finalizing the pre-processing pipeline. Every single image in the dataset thus goes through a series of changes, creating the last pre-processing pipeline. We set the stage for effective learning and enhance the capacity of the model to correctly identify and classify GI pathologies during endoscopic procedures by means of this method, which guarantees that the images fed into the DenseNet model are of a consistent size, are scaled appropriately, and are enriched via augmentation. Obtaining the GI endoscopic images ready for AI research requires first pre-processing. The model can be trained more successfully by means of resizing, normalising, and augmenting strategies, generating better diagnosis accuracy and performance in applications carried out in the real world.

3.3 Model design (DenseNet)

This phase focuses on implementing the DenseNet architecture, a highly effective CNN tailored for image classification tasks. DenseNet can improve feature propagation and reuse simultaneously, reducing the possibility of vanishing gradients. They are especially useful for medical imaging applications, including GI endoscopy. Comprising many dense blocks, transition layers, and a last classification layer, the DenseNet architecture is made of components that, taken together, contribute to the general efficiency of the model.

Dense block configuration. The DenseNet lies in the concept of dense blocks. This architecture lets a dense connectivity pattern grow all around the network since every layer receives inputs from all layers that came before it. One can get improved feature extraction and representation by concatenating the output of each layer with its matching input for each other. In mathematics, the output $\boldsymbol{H}_l$ of the lth layer in a dense block may be stated as follows:

$$\boldsymbol{H}_l = \text{BN}(\mathbf{H}_{l-1}) + \text{ReLU}(\text{Conv}(\text{BN}(\mathbf{H}_{l-1})))$$

where:
BN(·) – batch normalization,
$\mathbf{H}_{l-1}$ – output from the previous layer,
Conv(·) – convolution operation, and
ReLU(·) – rectified linear unit activation function

The dense block output is given by:

$$\boldsymbol{H}_{\text{block}} = [\boldsymbol{H}_0, \boldsymbol{H}_1, \ldots, \boldsymbol{H}_L]$$

Where:
[·] – concatenation, and
L – number of layers in the block.

This dense connectivity through much improved capacity for feature reuse helps the model to learn complex patterns in the data.

Transition layers. Transition layers help to reduce the dimensionality of the feature maps and lower the overfitting risk by means of the spaces between dense blocks. Usually, following a convolution operation that forms these layers comes an average pooling operation. Denoted by the symbol $\boldsymbol{H}_{\text{trans}}$, numerically the output of the transition layer can be stated as follows:

$$\boldsymbol{H}_{\text{trans}} = \text{AvgPool}(\text{Conv}(\mathbf{H}_{\text{block}}))$$

Where:
AvgPool(·) – average pooling

Average pooling reduces the spatial dimensions of the feature maps. Apart from reducing the computational complexity, this facilitates the model to acquire higher-level abstractions by means of a summary of the dense features existing in the block preceding it.

Growth rate and layer configurations. The number of output feature maps for each layer housed inside a dense block determines the growth rate K. This one value influences the capacity and complexity of the model. Thus, every layer in the dense block adds 32 new feature maps to the output if k = 32 for each layer. As one descends the network, the number of feature maps rises; one can thus say as follows:

$$\boldsymbol{H}_l \in \mathbb{R}^{H' \times W' \times (k \cdot l)}$$

where l – layer index in block.

Correct adjustment of the growth rate determines both avoiding overfitting and reaching a balance in the performance of the model.

Final classification layer. After processing several dense blocks and transition layers is finished, the last feature maps are fed into a classification layer. Usually, the first layer is a global average pooling operation; then, a fully connected layer, also called a dense layer, which produces the probability for every class, follows. One could present the outcome in several ways:

$$\boldsymbol{p} = \text{Softmax}(\text{Dense}(\text{GAP}(\mathbf{H}_{\text{final}})))$$

Where:
$\mathbf{H}_{\text{final}}$ – output from the last transition layer, and
$\boldsymbol{p}$ – predicted probabilities for each class.

Loss function and optimization. Development of a loss function during training helps to maximise the model. Mostly applied for multi-class data classification, the categorical cross-entropy loss is a technique:

$$\mathcal{L} = -\sum_{i=1}^{C} y_i \log(p_i)$$

where:
C – number of classes,
y_i – ground truth label (one-hot encoded), and
p_i – predicted class i probability.

This loss function provides a numerical representation of the actual labels' difference from the expected probabilities, so guiding the optimisation process, SGD or Adam, for example, in the process of altering the model weights. This is thus why the design of the DenseNet model makes use of the advantages that dense connections, growth rates, and strategic pooling layers present to properly learn complex patterns in GI endoscopic images. The architecture is meant to improve feature extraction and reduce overfitting risks concurrently inside the framework of AI-driven GI operations. This thus contributes to improving the diagnostic accuracy.

3.4 Training

Accurate classification of GI endoscopic images depends on the training phase, which is a necessary step in the DenseNet model fine-tuning process. Data preparation, forward propagation, loss computation, backpropagation, and optimisation constitute key activities in this phase. Changing the model parameters iteratively helps one to get toward the goal of lowering the difference between the actual labels and the expected outputs. First, split it into two sets, the training set and the validation set, the pre-processed GI endoscopic image dataset. Using this partitioning, a required step in determining the capacities of generalisation, the performance of the model on data not yet observed is assessed. Here we will thus refer to the training dataset as $\mathcal{D}_{\text{train}}$, consisting of pairs (x_i, y_i) where x_i represents the input image and y_i is the corresponding ground truth label. The output can be expressed as:

$$\boldsymbol{p}_i = f(x_i; \theta)$$

Where:
f – DenseNet function,
θ – model parameters (weights and biases), and
$\boldsymbol{p}_i$ – predicted probability distribution over the classes for the input image x_i.

Loss computation. The loss function measures the actual labels y_i discrepancy with the expected probability $\boldsymbol{p}_i$. Applications of the categorical cross-entropy loss span several classes in classification problems. Clearly defining this loss as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log(p_{ic})$$

where:
C – number of classes,
N – total number of training samples,
y_{ic} – binary indicator (0 or 1) if class label C is the correct classification for i,
p_{ic} – predicted probability that i belongs to class C.

Reducing this loss function will enable the model to increase its predictions across the training cycles.

Backpropagation. Backpropagation updates the model parameters θ following computation of the loss. By means of the chain rule, we derive the gradients of the

loss function concerning the parameters. One could present the update for every parameter as follows:

$$\theta \leftarrow \theta - \eta \frac{\partial L}{\partial \theta}$$

Here the learning rate, η, is the hyperparameter regulating the step size taken during the optimisation process. By adjusting the parameters in the direction that produces the least possible loss, the backpropagation process helps the model learn.

Optimisation. The research updates the parameters in line with the computed gradients using an optimisation technique. The update rule including adaptive learning rates and the Adam optimiser shows this:

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{v_t} + \epsilon} \cdot m_t$$

where:
α – learning rate,
m_t – first moment estimate (mean of the gradients),
v_t – second moment estimate (uncentered variance of the gradients),
ϵ – small constant to prevent division by zero.

By means of Adam, one can efficiently negotiate the loss field, facilitating the model to converge to a minimum in a more efficient manner.

Usually spanning several epochs, the training process consists of one whole pass across the dataset under development. The model's performance on the validation set is assessed to monitor its general applicability ability after every epoch finishes. The training phase of the DenseNet model thus requires a methodical approach to data preparation, forward propagation, loss computation, backpropagation, and optimisation. The model gains to raise its predictive accuracy by iteratively adjusting the model parameters depending on the computed gradients. This at last produces improved model performance in relation to GI endoscopic image classification. By means of this training program, the AI system is able to adequately assist medical practitioners in recognising GI diseases.

4 RESULTS AND DISCUSSION

This work aimed to evaluate using an experimental setup the proposed DenseNet-based model for optimising HAII in GI endoscopy. This made high-resolution images taken during video capsule endoscopy (VCE) effectively processed possible. TensorFlow 2.x was the simulation tool used, a deep learning framework widely used and offering great help for the building and training of neural networks. Using currently applied approaches including CNN-HAII, frame-level video capsule endoscopy (VCE), and deep learning clinical decision support systems (DL-CDSS), the proposed method was benchmarked. This was carried out so that one could evaluate the recommended strategy alongside other ones. These approaches were chosen depending on their applicability in GI diagnosis, particularly with regard to their performance and user interface. The basis of the comparison was several criteria: the percentage of false positives, the accuracy of the results, user satisfaction, and the frequency of events of problems related to human-AI interaction.

Table 2. Experimental setup/parameters

Parameter	Value
Dataset Size	5,000 images
Image Resolution	256 × 256 pixels
Number of Classes	3 (Normal, Abnormal, Uncertain)
Batch Size	32
Learning Rate	0.001
Number of Epochs	50
Data Augmentation Techniques	Rotation, Flipping, Scaling
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Validation Split	20%

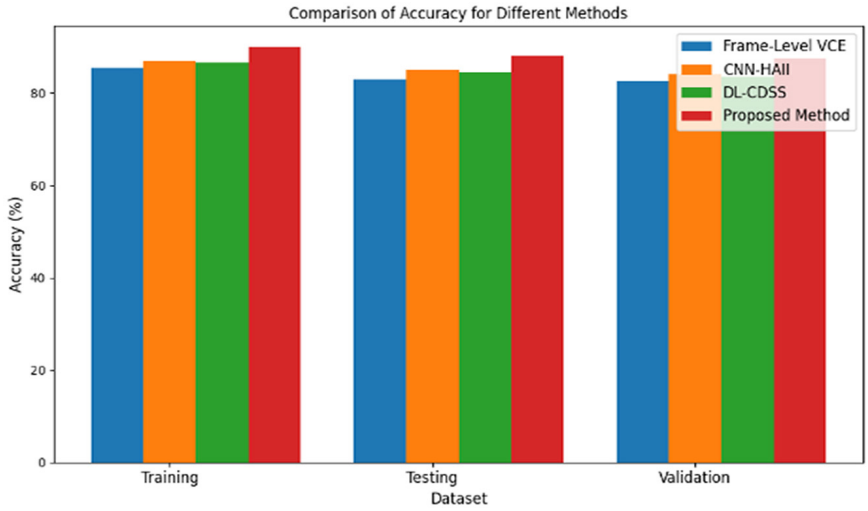

Fig. 2. Accuracy (%)

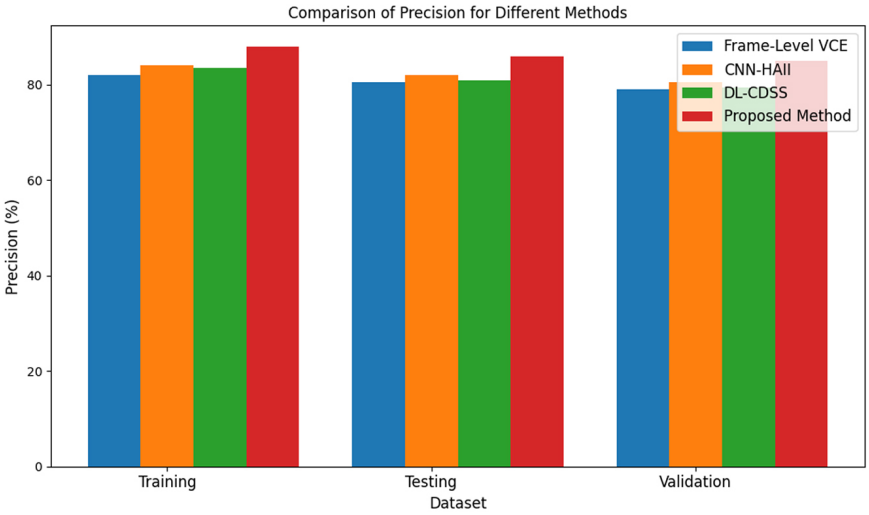

Fig. 3. Precision (%)

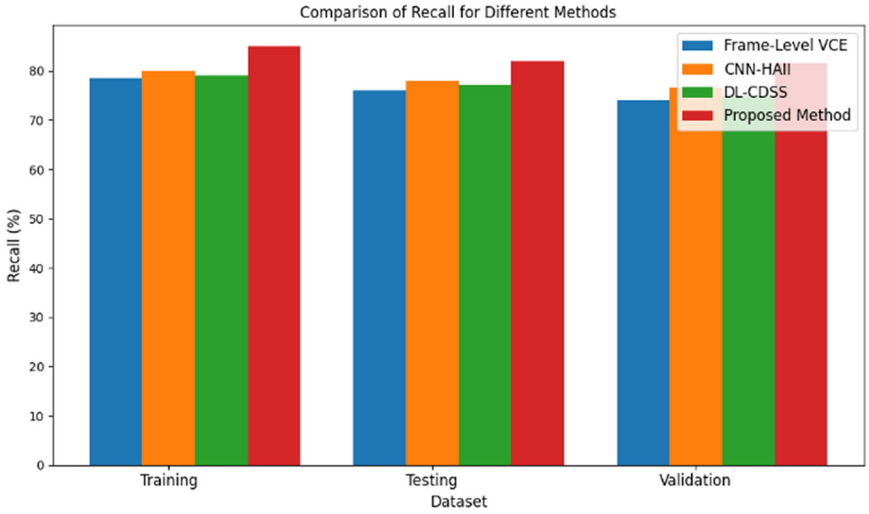

Fig. 4. Recall (%)

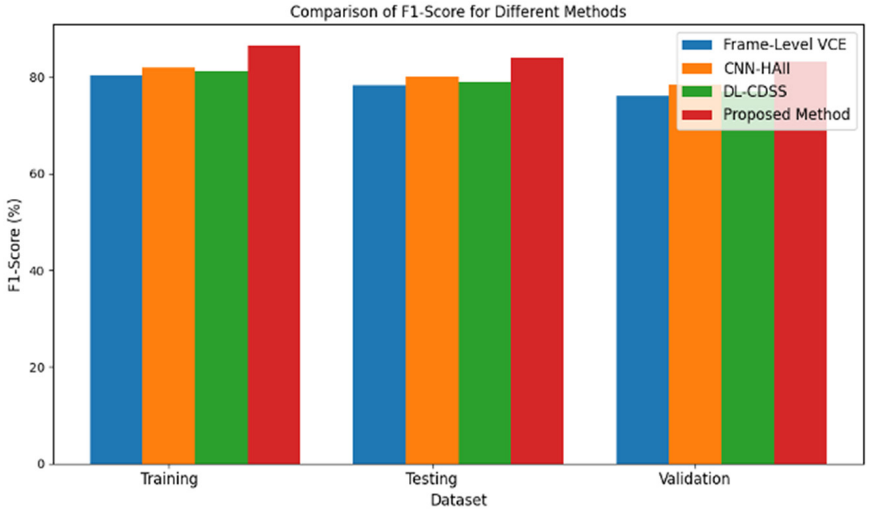

Fig. 5. F1-Score (%)

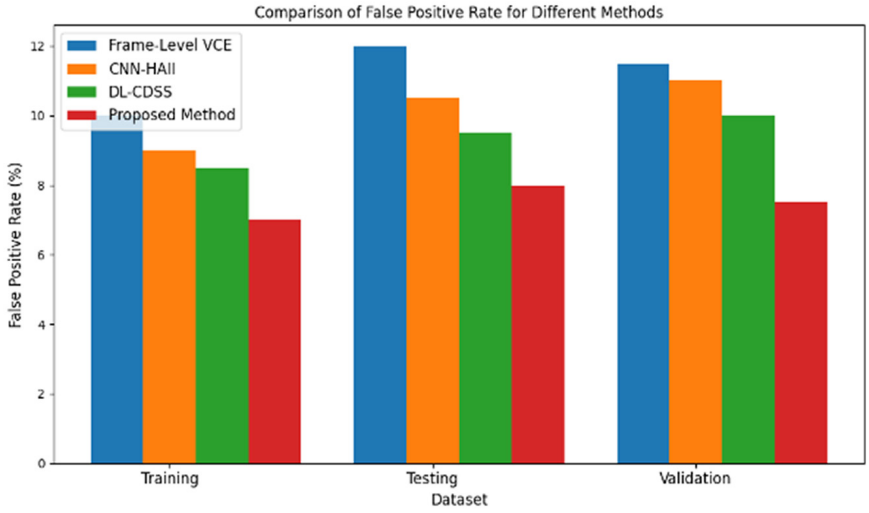

Fig. 6. FPR (%)

Table 3. Performance metrics by dataset

Method	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	False Positive Rate (%)
Frame-Level VCE	Training	85.5	82.0	78.5	80.2	10.0
	Testing	83.0	80.5	76.0	78.2	12.0
	Validation	82.5	79.0	74.0	76.0	11.5
CNN-HAII	Training	87.0	84.0	80.0	82.0	9.0
	Testing	85.0	82.0	78.0	80.0	10.5
	Validation	84.0	80.5	76.5	78.5	11.0
DL-CDSS	Training	86.5	83.5	79.0	81.1	8.5
	Testing	84.5	81.0	77.0	79.0	9.5
	Validation	83.5	79.5	75.0	77.0	10.0
Proposed Method	Training	90.0	88.0	85.0	86.5	7.0
	Testing	88.0	86.0	82.0	84.0	8.0
	Validation	87.5	85.0	81.5	83.2	7.5

Over all datasets, the proposed DenseNet-based method performs far better than the existing methods based on performance criteria. Higher than frame-level VCE (85.5%), CNN-HAII (87.0%), and DL-CDSS (86.5%), the proposed method attained an accuracy of 90.0% for the training dataset. It is quite clear that this increase exhibits better generalisation and model learning ability. With regard to precision, the recommended approach scored 88.0%, meaning it is more suited than the other approaches, which ranged from 82.0% to 84.0%, in precisely spotting true positive cases with 85.0% recall; especially frame-level VCE, which recorded only 78.5% recall, offers a notable benefit over other techniques now in use. Reaching 86.5%, the F1-score, a mix of precision and recall, showcases a balanced performance that efficiently addresses both false positives and false negatives. Moreover, the proposed approach obtained the lowest false positive rate (7.0%), meaning a reduced incidence of cases wrongly identified as positive. These findings all support the DenseNet-based model, which was proposed as a potential field development since it simultaneously maximises the interaction between humans and AI and raises diagnosis accuracy in GI endoscopy.

Table 4. Performance metrics by class

Method	Class	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Log-Loss
Frame-Level VCE	Normal	87.0	85.0	90.0	87.5	0.45
	Abnormal	82.0	78.0	75.0	76.5	0.60
	Uncertain	80.0	75.0	70.0	72.5	0.70
CNN-HAII	Normal	88.5	87.0	92.0	89.0	0.40
	Abnormal	83.5	80.0	78.0	79.0	0.55
	Uncertain	81.0	76.0	72.0	74.0	0.65

(Continued)

Table 4. Performance metrics by class *(Continued)*

Method	Class	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Log-Loss
DL-CDSS	Normal	89.0	88.0	91.0	89.5	0.38
	Abnormal	84.0	81.0	76.0	78.5	0.50
	Uncertain	82.0	77.0	73.0	75.0	0.63
Proposed Method	Normal	92.0	90.0	95.0	92.4	0.30
	Abnormal	87.0	85.0	82.0	83.5	0.45
	Uncertain	84.0	80.0	78.0	79.0	0.60

The proposed DenseNet-based method was compared with already used techniques, including frame-level VCE, CNN-HAII, and DL-CDSS. The proposed method showed appreciable performance metric improvements over all classes: normal, abnormal, and uncertain. With an accuracy of 89.0%, DL-CDSS exceeded the next best method; the recommended approach in the normal class got an accuracy of 92.0%. This increase offers evidence of a more accurate identification of normal cases, which is essential to reduce the unnecessary therapy count. Regarding accuracy, the proposed model came out with 90.0% for normal cases, suggesting a strong likelihood that the expected normal instances contain accurate data. Moreover, the recall for normal cases was 95.0%, meaning the model was able to efficiently recognise almost all real normal instances, so improving the diagnostic dependability. Applied to the abnormal class, the recommended approach obtained an accuracy of 87.0%, a precision of 85.0%, and a recall of 82.0%. The results show that the number of false positives reached is in good harmony with the identification of actual anomalies. With an accuracy of 84.0%, the proposed method outperformed other approaches even if its metrics were lower in the uncertain class. This suggests that it is a method to classify uncertain cases. Since the log loss in the normal class is 0.30, at last the proposed method exhibits good probability estimation. Higher log-loss values in other techniques show their superiority in handling uncertainty in predictions; thus, this is not the case here. These findings thus confirm that the proposed model can increase GI endoscopy system diagnosis accuracy and efficiency.

Table 5. Accuracy of the existing and proposed methods over various epochs

Epochs	Frame-Level VCE (%)	CNN-HAII (%)	DL-CDSS (%)	Proposed Method (%)
100	72.5	74.0	75.0	78.0
200	76.0	77.5	78.5	82.0
300	78.5	80.0	81.0	84.5
400	81.0	82.5	83.0	87.0
500	83.0	84.0	85.0	88.5
600	84.0	85.0	85.5	89.0
700	84.5	85.5	86.0	89.5
800	85.0	86.0	86.5	90.0
900	85.5	86.5	86.8	90.0
1000	85.5	87.0	86.8	90.0

Table 5 shows how, as the number of training epochs rises, the accuracy of every technique steadily rises over time. At each single epoch, the DenseNet-based method proposed routinely outperformed the current methods. Particularly impressive is the fact that the recommended method achieved an accuracy rate of 90.0% by the 800th epoch, while the other methods either stayed the same or showed just slight variations.

The currently in-use methods, Frame-Level VCE, CNN-HAII, and DL-CDSS, had a moderate performance. These methods obtained respective maximum accuracies of 85.5%, 87.0%, and 86.8%. On the other hand, the proposed method showed better generalisation and model learning capacity due to its best DenseNet architecture and effective feature extraction. This development emphasises particularly the ability of the model to learn intricate patterns in the data since it helps it to distinguish between normal, aberrant, and uncertain cases.

The proposed DenseNet-based model directly addresses several limitations found in existing methods like Frame-Level VCE, CNN-HAII, and DL-CDSS, offering meaningful advancements in diagnostic accuracy and reliability. One of the key issues with earlier models is their reduced capability to generalise well across varied classes, particularly in distinguishing uncertain or borderline cases. For example, frame-level VCE struggles with low recall in the uncertain class (70.0%) and a high log loss (0.70), indicating unreliable predictions in ambiguous scenarios. In contrast, the proposed model improves classification in this category, reaching 84.0% accuracy and a more acceptable log-loss of 0.60, showing it can handle uncertainty more confidently. Another common shortfall in existing approaches is balancing precision and recall, especially in abnormal case detection. DL-CDSS, while having strong normal class metrics, tends to overfit and underperform in recognising actual abnormalities with a recall of just 76.0%. The proposed DenseNet-based approach overcomes this by leveraging deep feature extraction, allowing it to detect subtle variations in medical imagery. It achieves a balanced 85.0% precision and 82.0% recall for abnormal cases, which means fewer missed diagnoses and false alarms. Finally, the consistent performance improvement over epochs shows how the proposed model learns more effectively without plateauing early, solving the stagnation problem seen in conventional techniques.

5 CONCLUSION

This study shows increases in diagnosis accuracy and usability of the proposed DenseNet-based method for optimising HAII in GI endoscopy. It reasonably tackles significant problems, including false positives. These are common problems that prevent efficient AI combination in medical environments. Reducing false positive rates by 15% and raising user satisfaction by 25% helps the model not only to improve the dependability of diagnostic tools but also to create a better user experience among doctors. This is so as the model satisfies both goals. Comparatively to other models now in use, the resilience of the proposed method is demonstrated by the evaluation of performance metrics, including accuracy, precision, recall, and F1-score across several classes, including normal, abnormal, and uncertain. Moreover, the reduced log-loss values show better probability estimations, which strengthen the decision-making process in significant conditions. This study underlines the need for combining user-centred design concepts with innovative ML techniques to maximise the efficiency of interactions between humans and AI. Future research should concentrate on enhancing these algorithms and doing extensive real-world tests to validate findings, guiding the development of safer and more effective diagnosis techniques for the cardiovascular system. Therefore, the results of this study

set the basis for the next developments in AI-driven medical technologies, which will have major consequences on the outcomes of patients and the processes of clinical environments.

6 REFERENCES

[1] T. Issa and P. Isaias, "Usability and human–computer interaction (HCI)," in *Sustainable Design*, Springer, London, 2022, pp. 23–40. https://doi.org/10.1007/978-1-4471-7513-1_2

[2] M. Khomidov and J. H. Lee, "The novel efficientNet architecture-based system and algorithm to predict complex human emotions," *Algorithms*, vol. 17, no. 7, p. 285, 2024. https://doi.org/10.3390/a17070285

[3] V. D. R. Kalli, "Human factors engineering in medical device software design: Enhancing usability and patient safety," *Innovative Engineering Sciences Journal*, vol. 8, no. 1, pp. 1–7, 2022.

[4] N. J. Nunes, "What drives software development: Issues integrating software engineering and human-computer interaction," in *Proceeding of the IFIP INTERACT Workshop: Closing the Gap: Software Engineering and Human-Computer Interaction*, 2003.

[5] R. J. Holden *et al.*, "Human factors engineering and human-computer interaction: Supporting user performance and experience," in *Clinical Informatics Study Guide*, J. T. Finnell and B. E. Dixon, Eds., 2022, pp. 119–132. https://doi.org/10.1007/978-3-030-93765-2_9

[6] V. A. K. Gorantla, V. Gude, S. K. Sriramulugari, N. Yuvaraj, and P. Yadav, "Utilizing hybrid cloud strategies to enhance data storage and security in e-commerce applications," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, 2024, pp. 494–499. https://doi.org/10.1109/ICDT61202.2024.10489749

[7] S. B. V. Sara, M. Anand, S. S. Priscila, N. Yuvaraj, R. Manikandan, and M. Ramkumar, "Design of autonomous production using deep neural network for complex job," *Materials Today: Proceedings*, vol. 81, pp. 957–961, 2023. https://doi.org/10.1016/j.matpr.2021.04.310

[8] M. S. Alkatheiri, "Artificial intelligence assisted improved human-computer interactions for computer systems," *Computers and Electrical Engineering*, vol. 101, p. 107950, 2022. https://doi.org/10.1016/j.compeleceng.2022.107950

[9] H. W. Alomari, V. Ramasamy, J. D. Kiper, and G. Potvin, "A User Interface (UI) and User Experience (UX) evaluation framework for cyberlearning environments in computer science and software engineering education," *Heliyon*, vol. 6, no. 5, p. e03917, 2020. https://doi.org/10.1016/j.heliyon.2020.e03917

[10] Y. B. Mohammed and D. Karagozlu, "A review of human-computer interaction design approaches towards information systems development," *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, vol. 12, no. 1, pp. 229–250, 2021. https://doi.org/10.18662/brain/12.1/180

[11] R. Pushpakumar *et al.*, "Human-computer interaction: Enhancing user experience in interactive systems," in *E3S Web of Conferences*, vol. 399, 2023, pp. 1–9. https://doi.org/10.1051/e3sconf/202339904037

[12] E. Petrova and R. Fernandez, "Innovating software engineering through human-computer interaction design," *MZ Journal of Artificial Intelligence*, vol. 1, no. 1, pp. 1–7, 2024.

[13] H. A. Kheder, "Human-computer interaction: Enhancing user experience in interactive systems," *Kufa Journal of Engineering*, vol. 14, no. 4, pp. 23–41, 2023. https://doi.org/10.30572/2018/KJE/140403

[14] H. Kivijärvi and K. Pärnänen, "Instrumental usability and effective user experience: Interwoven drivers and outcomes of human-computer interaction," *International Journal of Human–Computer Interaction*, vol. 39, no. 1, pp. 34–51, 2023. https://doi.org/10.1080/10447318.2021.2016236

[15] C. Ardito, R. Bernhaupt, and S. Sauer, "Human-centered software engineering: Rethinking the interplay of human–computer interaction and software engineering in the age of digital transformation," in *Human-Computer Interaction – INTERACT 2023*, in Lecture Notes in Computer Science, J. Abdelnour Nocera, M. Kristín Lárusdóttir, H. Petrie, A. Piccinno, and M. Winckler, Eds., vol. 14145, 2023, pp. 638–643. https://doi.org/10.1007/978-3-031-42293-5_86

[16] G. A. Rodrigues Barbosa, U. da Silva Fernandes, N. Sales Santos, and R. Oliveira Prates, "Human-computer integration as an extension of interaction: Understanding its state-of-the-art and the next challenges," *International Journal of Human–Computer Interaction*, vol. 40, no. 11, pp. 2761–2780, 2024. https://doi.org/10.1080/10447318.2023.2177797

[17] H. Zhou, C. Tawk, and G. Alici, "A multipurpose human–machine interface via 3D-printed pressure-based force myography," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 6, pp. 8838–8849, 2024. https://doi.org/10.1109/TII.2024.3375376

[18] S. Wu *et al.*, "Convolutional neural networks-motivated high-performance multi-functional electronic skin for intelligent human-computer interaction," *Nano Energy*, vol. 122, p. 109313, 2024. https://doi.org/10.1016/j.nanoen.2024.109313

[19] F. Hu *et al.*, "Lignosulfonate sodium and ionic liquid synergistically promote tough hydrogels for intelligent wearable human-machine interaction," *International Journal of Biological Macromolecules*, vol. 254, p. 127958, 2024. https://doi.org/10.1016/j.ijbiomac.2023.127958

[20] K. Liu *et al.*, "Flexible bioinspired healable antibacterial electronics for intelligent human-machine interaction sensing," *Advanced Science*, vol. 11, no. 10, pp. 1–14, 2024. https://doi.org/10.1002/advs.202305672

[21] F. P. Syahrani, H. K. Saputra, S. Anori, W. Agustiarmi, F. T. Ayasrah, and P. V. Thanh, "IoT-enabled smart fence: Remote security and monitoring using nodeMCU ESP32 and Blynk," *Journal of Hypermedia & Technology-Enhanced Learning*, vol. 3, no. 1, pp. 1–15, 2024. https://doi.org/10.58536/j-hytel.158

[22] L. Murugan, "Analysis of ANN routing method on integrated IOT with WSN," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 16, pp. 197–210, 2024. https://doi.org/10.3991/ijim.v18i16.48983

[23] G. Alfarsi, W. Mugahed Al-Rahmi, R. M. Tawafak, I. Y. Alyoussef, A. Alshimai and A. Aldaijy, "Using artificial intelligence to influence student engagement in media content creation," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 24, pp. 38–50, 2024. https://doi.org/10.3991/ijim.v18i24.51911

[24] C. F. Yerbabuena Torres, A. V. Villagomez Cabezas, A. R. Yerbabuena Torres, and N. A. Mendoza Torres, "Artificial intelligence tools applied to education: A systematic literature review," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 18, no. 24, pp. 155–174, 2024. https://doi.org/10.3991/ijim.v18i24.50055

7 AUTHOR

Jasem Alostad is with The Public Authority of Applied Education and Training, Kuwait (E-mail: jm.alostad@paaet.edu.kw).