Editorial

We present the third issue of the sixteenth volume of the **Journal of Multimedia Processing and Technologies** with the research described as follows.

In the opening paper, ***"Temporal Dynamics and Latent Structures of Emotional States in Human–Robot Interaction: A Markovian and Hidden Markov Model Analysis of Longitudinal Facial Emotion Sequences,"*** the authors argue that grasping the temporal progression of human emotions is essential in Human Robot Interaction (HRI) systems. To explore long-term dynamics and uncover hidden affective patterns, they introduce the Temporal Emotion Dynamics and Latent Regime Framework (TEDLRF). This framework models and analyzes longitudinal sequences of five core emotions using Markov Chains, Hidden Markov Models (HMMs), and Dynamic Bayesian Networks. Their findings suggest that observable facial expressions reflect deeper, unobservable psychological states. As a result, they contend that next-generation adaptive robots must evolve beyond reactive, imitation-based behaviors toward proactive, temporally aware system designs.

The subsequent paper, ***"Automated Skin Lesion Detection and Instance Segmentation Using Mask R-CNN: A Comprehensive Pipeline Analysis,"*** presents a complete end-to-end computational pipeline for automated skin lesion detection and instance segmentation based on the Mask R-CNN architecture. The work aims to develop a robust system capable of simultaneous lesion localization, classification, and pixel-level segmentation, while also delivering clinically interpretable outputs to aid dermatological decision-making. Their proposed framework incorporates a ResNet-50 backbone with a Feature Pyramid Network (FPN) for multiscale feature extraction, a Region Proposal Network (RPN) for candidate detection, ROI Align for precise spatial feature mapping, and parallel multitask heads for classification, bounding-box regression, and mask prediction. The system achieves strong performance on validation data, underscoring its potential as a reliable, accurate, and clinically applicable tool for automated skin lesion analysis.

The final paper, ***"Quantitative Evaluation of Autoencoder Reconstruction Performance on the CIFAR-10 Dataset,"*** centers on autoencoders foundational unsupervised neural networks that learn compact data representations by minimizing reconstruction error. The author conducts a comprehensive quantitative and qualitative assessment of a convolutional autoencoder (CAE) applied to the CIFAR-10 dataset. A multidimensional evaluation framework is employed, combining pixel-level metrics (MAE, MSE, RMSE), signal quality (PSNR), and perceptual similarity (SSIM), complemented by diagnostic visualizations such as side-by-side comparisons, error heatmaps, per-image error distributions, and class-wise performance breakdowns. Results indicate that semantically meaningful class structures emerge without explicit supervision, with confusion heatmaps showing 89–95% diagonal alignment with ground truth labels.

Altogether, the research featured in this issue exemplifies both elegance and efficacy, offering powerful solutions to a range of complex challenges in multimedia technologies.

Editors