HybridFusionNet: Hybrid Fusion Network for Underwater Image Restoration

Anshika
Department of Computer Science and Engineering
IIT Ropar
Rupnagar, Punjab
2021csb1069@iitrpr.ac.in

***Abstract*—Water causes several adverse effects on underwater images such as color distortion, visibility issues, wave generated noise. This paper presents a novel hybrid fusion method that integrates deep learning based convolution neural networks (CNN) with advanced fusion strategies. The proposed model extracts spatial, temporal, and color features through a combination of early and late fusion techniques. Specifically, stacked autoencoders (SAEs) are employed for spatial feature extraction, while long short-term memory (LSTM) networks are utilized for temporal feature learning. A decision-level fusion mechanism is used to integrate these features for achieving enhanced restoration accuracy. Experimental evaluations demonstrate that the proposed method significantly improves the quality of restored underwater images. This approach offers a robust and efficient solution for real-time underwater vision applications in areas such as marine biology, underwater robotics, and environmental monitoring. The code implementation can be accessed from https://github.com/anshikabhatia13/HybridfusionNet**

***Index Terms*—Hybrid Fusion, Fusion Models, Underwater Image Restoration, Convolutional Neural Network, Long Short-Term Memory**

I. Introduction

Underwater image restoration remains a critical challenge in computer vision due to the special optical distortions of the underwater environment such as light absorption scattering and reduced contrast [1]. These special optical distortions reduce image quality significantly, making it more difficult for underwater vision systems to perform tasks like mapping, object detection, and classification. Due to the intricate interactions between underwater light and water characteristics, traditional restoration techniques frequently have difficulty in effectively denoising and enhancing the images. In deep learning, the recent advancements in the fields of convolutional neural networks (CNNs) and, crucially, hybrid fusion models have shown immense potential. Hybrid fusion models combine several data types and modelling techniques, are developed for the growing need for predictive systems that are accurate, effective, and interpretable and can handle multiple aspects of the data [2]. Through the use of data- driven representations, domain-specific knowledge, and the synergistic benefits of early and late fusion, these models aim to overcome the drawbacks of solitary architectures. Traditional approaches frequently fall short of capturing complex connections across geographical, temporal and feature dimensions, particularly in dynamic, high-dimensional contexts like image processing, traffic flow predictions, and weather modeling. This makes it important to design the hybrid architectures that are resilient, scalable, and noise-resistant and they can effectively fuse heterogeneous inputs but also maintain the robustness, scalability, and noise resilience .

In the work [3], Jin et al. introduces a hybrid spatiotemporal traffic prediction model which combines Stacked Autoencoders (SAEs) and Long Short-Term Memory (LSTM) networks. This is motivated by independently extracting spatial features from upstream toll gate data using SAEs and temporal dependencies from previous flow data using LSTM. The model therefore, incorporates these representations at a merging layer prior to predicting regression, exemplifying that deep fusion improves accuracy and capture cross-domain linkages in such data. Similarly, Feral et al. [4] presented HYCELL, a new simulation framework for rain-rate fields in environmental modelling. Their hybrid approach creates realistic rainfall fields for satellite and wireless applications by repeatedly modifying Gaussian and exponential rain-cell components to fit local cumulative distribution functions. The possibility of combining statistical learning with domain knowledge is highlighted by HYCELL's hybrid nature, which combines physical modelling with statistics.

In this work, we introduce a novel underwater image restoration model HybridFusionNet that integrates deep learning with hybrid fusion techniques. The proposed model combines spatial, temporal, and color feature extraction of the temporally related underwater dataset, using deep networks. The fusion of their outputs is performed at early and late stages. With the help of CNNs, the model is able to capture hierarchical features for the restoration. The fusion is further enhanced by the incorporation of stacked autoencoders (SAEs) for spatial features and long short-term memory (LSTM) networks for the temporal dependencies. Finally, a decision-level fusion mechanism is used to combines network outputs and improve the accuracy and robustness of the restoration, demonstrating the significance of hybrid models in underwater image restoration.

II. Related Work

In this section we outline the significance of hybrid fusion architectures in various domains,and highlight their ability for enhancing robustness, accuracy, and adaptability by integrating diverse fusion strategies and modeling paradigms Hybrid fusion architectures have gained great prominence across diverse domains such as fake news detection, soft sensing in industrial processes, health prognostics, and fuzzy classification systems. In these systems the strengths of different fusion strategies like early and late fusion, data-level and decision-level fusion, knowledge-driven and data-driven modeling are combined to enhance the robustness, accuracy, and adaptability of the mdoels.

Hamed et al. [5] propose HF-TIM, a hybrid fusion approach which integrates early (Level-0) and late fusion (Level-1) techniques for the fake news detection application. Softmax classifier performs early fusion to process the multimodal data. Here, the BERT base model is used to extract the textual information and VGG-19 model brings out image features. The image and text features are concatenated and processed through multiple layers and generate multimodal predictions. Along with multimodal predictions, independent predictions of textual features (BERT) and image features (VGG-19) are made using unimodal classifiers to ensure the capturing of distinct features independently, before fusion as well. This process is followed by Level-1 late fusion with a meta-learning classifier which uses the initial predictions of BERT, VGG-19 and Softmax, and integrates them to generate a final classification. The late-fusion system captures different but related dependencies as it considers unimodal and multimodal relationships using stacking methods.

Wang et al. [6] propose a weight-fusion and model risk assessment-based soft sensor for online prediction of primary variables for batch processes. The hybrid model, based on data-driven and mechanism modeling primarily uses parameter optimization-based or error-compensation-based structures. While the parameter optimization-based method uses data from the process to optimize the mechanism model with unknown parameters by minimizing the RMSE, the latter avoids complex optimization by using offline input-output data to correct errors. The parameter optimization-based method suffers from getting stuck at local minima, while the error-compensation-based method is sensitive to noise and may overfit. The proposed model utilizes a weighted sum of the predictions by both data-driven and mechanism models such that the mechanism model plays the primary role and its errors are then compensated by the data-driven model (error-compensation-based structure). To further enhance the robustness of the proposed model, accurate weights are determined by the risk assessment method which is based on Bayesian Information Criterion (BIC) and Structural Risk Minimization (SRM).

Ramani et al. [7] propose a hybrid data classification model that combines deep learning with an Adaptive Lion Fuzzy System (ALFS) and a Robust Grey Wolf-based Sine Cosine Algorithm (RGSCA-FS). ALFS applies fuzzy rules for its classification while optimized fuzzy rules generated by RGSCA-FS and Grey Wolf Optimization(GWO) are used for their classification output. Convolutional neural networks (CNN), identify complex patterns using multiple layers. Results of all three ALFS, RGSCA-FS and DCNN are fused for the hybrid classification output. The proposed ALFS makes classification rules from training data by applying fuzzy rules where prime components are fuzzification, fuzzy inference and defuzzification. The input data undergoes fuzzification to give fuzzy values, followed by fuzzy inference where these values are compared with fuzzy rule base. In fuzzification, the input (typically scalar) gets mapped to the fuzzy range of [0,1] using membership functions. Fuzzy inference matches fuzzy value of input with if-then rules set. The set of rules is generated using the Adaptive Lion Algorithm (ALA) and is associated with a weight indicating its strength. Defuzzification converts fuzzy decision to give class labels and the final output of classification is based on the maximum degree of membership of a class. This model achieves a classification accuracy of 94.11%, outperforming traditional fuzzy and data mining approaches.

Wen et al. [9] propose a dual-level fusion model for remaining useful life (RUL) estimation in industrial predictive maintenance scenarios. The architecture integrates data-level fusion (using Genetic Programming to generate compound health indicators from sensor data) and decision-level fusion (using Belief Function Theory to combine uncertainty-aware predictions). The model deploys computation across a three-plane architecture: equipment (local, low-latency analytics), edge (real-time inference), and cloud (model training and optimization). Advanced fusion strategies such as AdaBoost and Stacking are used alongside traditional statistical methods. Dempster's Rule is applied in the belief function framework to merge uncertainty-laden predictions from multiple sensor sources, thereby increasing the reliability and precision of RUL estimates under noisy and decentralized IIoT conditions.

Wei et al. [10] propose a hybrid dual-level fusion architecture for estimating remaining useful life in industrial applications for predictive maintenance. The proposed system is a combination of data-level and decision-level fusion for creating a robust health management system for IIoT environments. IIoT environments have a wide variety of sensors such as pressure, vibration, temperature and many more. This makes the task of predicting whether an equipment is likely to fail or not, i.e. its remaining useful life (RUL). In the data level fusion, data from multiple sources is integrated to create a compound indicator for health, whereas in the decision level fusion, the RUL estimates are combined. The system manages real-time, complex data coming from multiple sources by using Genetic Programming (GP) for heuristic data-level fusion and Belief Functions Theory (BFT) for decision-level uncertainty management. The fusion is done using AdaBoost and Stacking for the estimation of remaining useful life of IIoT systems. Due to the large scale of sensor data across a factory, a centralized predictive system is not viable. Hence,

TABLE I
SUMMARY OF HYBRID FUSION MODELS FROM LITERATURE

Model	Domain	Fusion Type	Techniques	Strengths	Limitations
Jin et al. [3]	Traffic Prediction	Spatiotemporal	SAEs (spatial), LSTM (temporal), Merge + ReLU layers	Space-time capture, Robust and scalable	Needs synchronized data, Location-dependent
Feral et al. [4]	Meteorology	Statistical Hybrid	HYCELL (Gauss+Exp), Iterative field simulation	Realistic rain fields, Geographically flexible	Limited to rain, Sensitive to climate input
Zang et al [8].	Image Fusion	Attention-Based	CNN + UFA, Hybrid loss (L1 + SSIM)	Detail preservation, Artifact-free fusion	High computational cost, Needs diverse training data
Wang et al. [6]	Medical Imaging	Layered Fusion	VM + SR + WSpNM, ADMM solver	Sharp detail retention, Robust to noise	Slower convergence, Complex tuning
Ramani et al. [7]	Healthcare IoT	Early-Late Hybrid	Early fusion of sensors, Late fusion via decision layer	Reduces uncertainty, Real-time feedback	May overfit with sparse data, Sync issues in early fusion
Wen et al. [9]	Remote Sensing	Cross-modal Fusion	Transformer + CNN, Feature alignment module	Handles modality gap, Improves semantic fusion	Transformer overhead, Needs large memory
Hamed et al. [5]	Risk Forecasting	Risk-aware Hybrid	Multi-stream fusion, Fuzzy entropy weighting	High interpretability, Adaptive risk modeling	Complexity in tuning, Slower in dynamic setups

in the proposed system, computation is distributed across three layers, the equipment plane, the edge plane and the cloud plane. The hybrid model reduces sensor noise, handles uncertainty and optimizes maintenance.

Jin et al. [3] propose a hybrid model for traffic forecasting where the data coming from five upstream spatial toll collection gates and one remote microwave sensor (RTMS) is combined. The spatial and temporal features of the data are also combined by the hybrid model to predict the traffic flow with accuracy. Deep learning techniques such as Stacked Autoencoders (SAEs) and Long Short-Term Memory (LSTM) are used to capture spatial and temporal features. The SAEs are used to extract spatial features from spatial data (upstream toll gate entries) while LSTM networks obtain the temporal features (flow of traffic at previous time steps). The toll gate data simply contains the entry and exit location-time of vehicles giving insight into the volume and origin of the traffic while the RTMS and Traffic video data (using Traffic Video Detection Equipment) give temporal information in detail in real-time. The SAEs-LSTM model brings average reduction in mean absolute error and root mean squared error by 6-15%.

Feral et al. [4] propose a novel method for simulating rain-rate fields. The designed method is for two-dimensional rain rate fields, which are spread over 10s to 100s sq. kilometers. This is applicable for the precipitation impact on the telecommunication beams of satellites and in the terrestrial wireless access networks of broadband. In the proposed approach, the rain scenarios are generated consistently with the local climate and its characteristics over the area where the simulation is performed. The central component of this methodology is the usage of the HYCELL model for representing single rain cells along with analytical expression of their spatial densities. The statistical distribution of the rain cell is the widely studied parameter from which the analytical expressions originate. The local cumulative distribution function of the rate of the rain is taken as input, generating the rain fields which represent the meteorological conditions locally. This local information, along with sophisticated modelling, shows great improvement in the overall accuracy of meteorological simulation.

Wang et al. [6] propose Hybrid Variation-Sparse Representation (HSVR) for denoising and fusion taks in medical images. The proposed architecture is a combination of three mathematical concepts: the Variation Model (VM), Sparse Representation (SR), and the Weighted Schatten p-Norm Minimisation (WSpNM). VM gives a noise-free base layer. This is done by the removal of texture noise and the extraction of structure in the noisy image by local mean isolation. The SR module helps in getting continuous texture details from the noisy inputs. However, it can lead to loss of information because of sparsity constraints. WSpNM applies weighted penalties to the singular values. This leads to enhanced performance in denoising by gradient detail preservation. The VM, SR and WSpNM give three primary components from the inputs. HSVR is not only a decomposition system but also an advanced denoising system. The first stage is decomposition. The second stage is the application of the fusion rules to combine these outputs. Then finally the features are fused selectively, constructing the detail layer which uses a weighted sum over all inputs.

Li et al. [11] propose a GPS/INS system which integrates the data from a MEMS-based IMU and the GPS. This proposed system is structured into two phases: the training phase and the prediction phase. The raw acceleration and the gyroscope readings are provided by the IMU along the three axes x,y,z. The training phase focuses on learning INS error characteristics from the available GPS data. Secondly, IMU signals are taken into account such that they are denoised using the Empirical Mode Decomposition with Wavelet Denoising (EMD-WD) module and this denoised data is used to calculate the INS-estimated position using mechanisation equations. An Interactive Multi-Model Extended Kalman Filter is used to output the correct position by using the difference between GPS and INS positions as input. Using the corrected output of the Kalman filter as a label, an Extreme Learning Machine (ELM) is trained. This ELM takes the INS position and the time T as the input. The prediction error during training is calculated as the difference between the Kalman corrected position and the predicted position by ELM. The hybrid fusion methods are

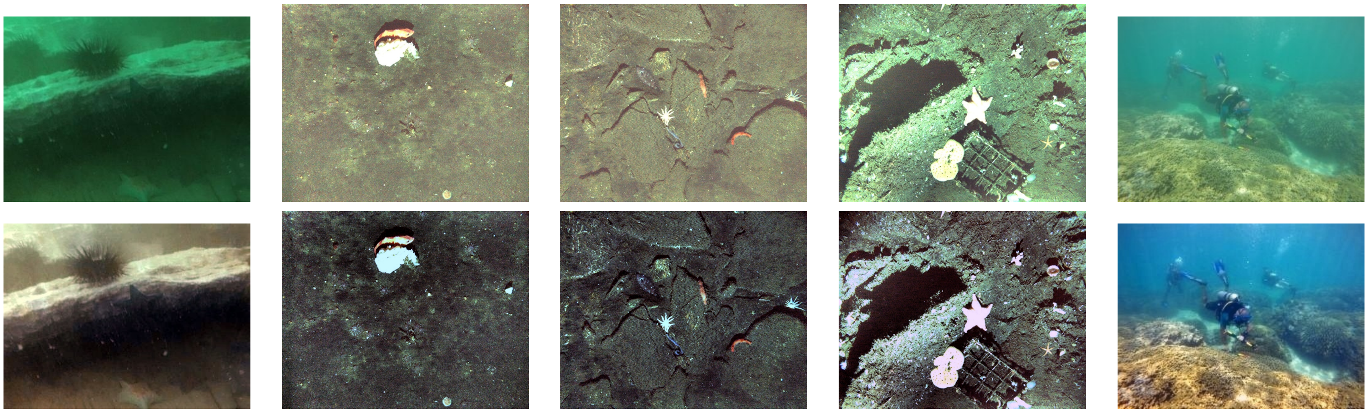

Fig. 1. Top row: Degraded images. Bottom row: Corresponding restored images using Proposed HybridFusionNet

also gaining popularity in the domain of underwater image restoration and enhancement [12] propose an unsupervised image restoration model for degenerated underwater images. They construct this model using the CycleGAN architecture and explicit physical inferences based on the degeneration model by Jaffe-McGlamery . The scene depth, backscatter factors, attenuation factors, and veiling light make up degenerated underwater photographs, which are separated into content and style. This method gives color-corrected image by rearranging these separate components, which improves restoration performance over a single end-to-end generator. The improved performance of the suggested hybrid physical-neural solution is demonstrated along with proof that the backscatter loss and the concatenated input of the degeneration factor networks are powerful cues for improving depth estimation, global contrast, and color correction, all of which artificially aid in restoration. Similarly, the effectiveness of various levels hybrid fusion for underwater image restoration is demonstrated in [2], [13]. In [1], local fusion is used for image enhancement in underwater data. An adaptive transmission fusion strategy is used in [14] for the same application.

In this paper we propose HybridFusionNet, a novel hybrid fusion network specifically designed for underwater image restoration

III. Proposed HybridFusionNet for Underwater Image Restoration

In this section we introduce the proposed HybridFusionNet, a novel deep learning-based hybrid fusion model designed for the task of high-fidelity restoration of underwater images. The proposed hybrid fusion network architecture integrates spatial encoding, temporal modeling, attention-driven feature fusion, and perceptual enhancement. The output restored images are visually and structurally enhanced. The model consists of the following components which result in the hybrid fusion model

A. Stacked Autoencoder

The stacked autoencoder component of the overall model is employed to extract the hierarchical spatial features from each underwater image frame. During decoding,to preserve structural details, this module uses skip connections with refinement layers. The encoder-decoder network can be represented as follows: $\mathbf{x}_t \in \mathbb{R}^{H\times W\times C}$:

$$\mathbf{z}_t = f_{\text{enc}}(\mathbf{x}_t), \quad \hat{\mathbf{x}}_t = f_{\text{dec}}(\mathbf{z}_t)$$

where f_{enc} is the encoder and f_{dec} is the decoder.

B. Temporal LSTM:

The Temporal LSTM module helps to capture the temporal coherence across multiple frames in a sequence. It also extracts the global temporal trends and encodes them to guide image restoration.

$$\mathbf{h}_t, \mathbf{c}_t = \text{LSTM}(\mathbf{z}_t, \mathbf{h}_{t-1}, \mathbf{c}_{t-1})$$

where $\mathbf{h}_t$ and $\mathbf{c}_t$ are the hidden and cell states, respectively.

C. Multiscale Attention Fusion Module

The Multiscale Attention Fusion Module incorporates the spatial and temporal attention. This is done for an adaptive feature selection. Firstly, the multi-scale convolutions are used to capture all kind of features from fine to coarse image features. This is followed by the progressive upsampling of the hybrid fused components to the original resolution. The multiscale features are combined and attention weights α_i are computed as follows:

$$\alpha_i = \frac{\exp(g(\mathbf{f}_i))}{\sum_{j=1}^{N}\exp(g(\mathbf{f}_j))}, \quad \mathbf{F}_{\text{fused}} = \sum_{i=1}^{N}\alpha_i \cdot \mathbf{f}_i$$

where $\mathbf{f}_i$ is the feature map at scale i and $g(\cdot)$ is a learnable scoring function.

D. Color Refinement Network and Hybrid Loss Function

The Color Refinement Network in the model is there to enhance the chromatic and textural fidelity. For this, residual-style convolutional blocks are used. The Hybrid Loss Function is a combination of MSE (reconstruction), SSIM (structure), and VGG-based perceptual loss. The hybrod loss function is used to ensure a perceptual quality and structural similarity in the restored underwater outputs.The total loss function

TABLE II
COMPARISON OF OUR HYBRID FUSION MODEL WITH STATE-OF-THE-ART METHODS

Model	PSNR (dB)	Parameters (M)	Year
Proposed HybridFusionNet*	26.99	2.8	2025
UGAN [15]	22.40	3.2	2018
Water-Net [16]	23.82	4.5	2020
DUIENet [17]	24.15	5.1	2020
FUnIE-GAN [18]	22.86	3.3	2021

combines reconstruction, perceptual, and structural similarity terms which can be defined as follows:

$$\mathcal{L}_{\text{total}} = \lambda_{\text{rec}}\mathcal{L}_{\text{rec}} + \lambda_{\text{ssim}}\mathcal{L}_{\text{SSIM}} + \lambda_{\text{perc}}\mathcal{L}_{\text{VGG}}$$

Reconstruction Loss (MSE):

$$\mathcal{L}_{\text{rec}} = \frac{1}{N}\sum_{i=1}^{N} \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2^2 \quad (1)$$

Structural Similarity Loss:

$$\mathcal{L}_{\text{SSIM}} = 1 - \text{SSIM}(\hat{\mathbf{x}}, \mathbf{x}) \quad (2)$$

Perceptual Loss (VGG):

$$\mathcal{L}_{\text{VGG}} = \sum_{l} \|\phi_l(\hat{\mathbf{x}}) - \phi_l(\mathbf{x})\|_2^2 \quad (3)$$

where $\phi_l(\cdot)$ is the activation from layer l of a pretrained VGG-19 network.

IV. EXPERIMENTAL SETTING

To evaluate the effectiveness of our proposed hybrid fusion model for underwater image restoration, we conducted extensive experiments under a controlled and reproducible setting. The experiments were implemented using PyTorch and conducted on a high-performance GPU-enabled system.

A. Dataset Preparation

The proposed HybridFusionNet is trained on the UIEB dataset. This is a comprehensive underwater image enhancement dataset with includes large-scale real-world images.The Underwater Image Enhancement Benchmark (UIEB) includes 950 real-world underwater images, 890 having the corresponding reference images. The dataset is split into training and testing splits in 7:3 in which each training sample is a sequence of 3 consecutive degraded frames, and the last frame corresponds temporally to the target ground truth image.

B. Preprocessing and Augmentation

The dataset is preprocessed and augmented using several transformations to enhance the model generalizability and robustness during both training and validation. The preprocessing steps during the training include training augmentations where the random horizontal and vertical flipping on input images is done. For better generalizability over different colours, color jitterinf is done with the adjustments in brightness, contrast and saturation. Further to enhance rotational robustness, random rotations on the data is performed. All the input images are resized to256×256 during both training and validation preprocessing. The normalization is also performed with mean and standard deviation of (0.5, 0.5, 0.5) during both training and validation preprocessing.

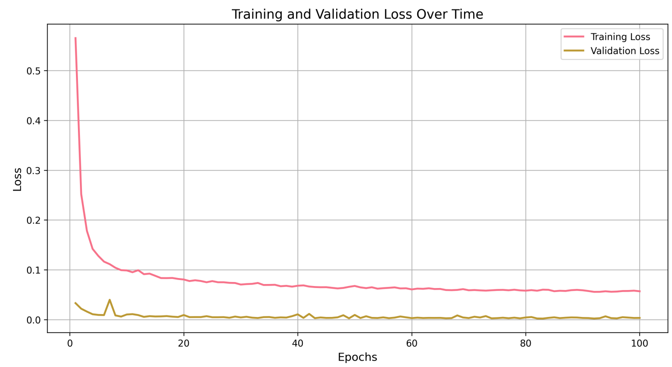

Fig. 2. Loss Curves

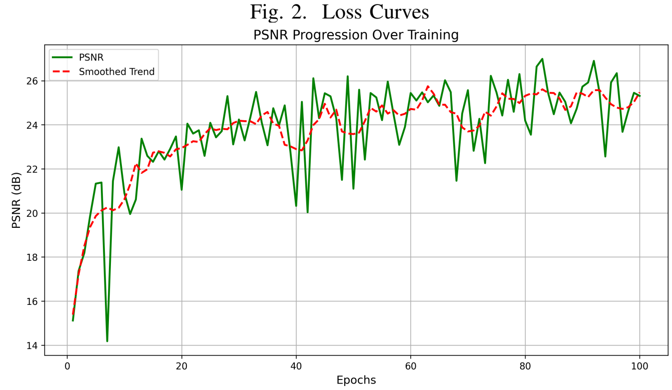

Fig. 3. PSNR vs Epochs

C. Model Configuration and Optimization Settings

In the proposed HybridFusionNet, the hybrid fusion network integrated three components: a stacked autoencoder for spatial encoding, an LSTM for temporal modeling, and an attention-guided decoder. The model configuration includes three input channels with a sequence length of 3. The loss is a composite hybrid with underlying functions defined in **equation ??**. The L1 loss has a weightage of 0.5, MSE loss has a weightage of 0.2, and VGG-based perceptual loss is given a weight of 0.3 for hybrid fusion.

For optimization of the proposed HybridFusionNet, AdamW optimizer is used. The learning rate is set to 1e-4, with weight decay as 1e-4 as well. The batch size of 16 is used and the model is trained for 100 epochs. For preventing overfitting due

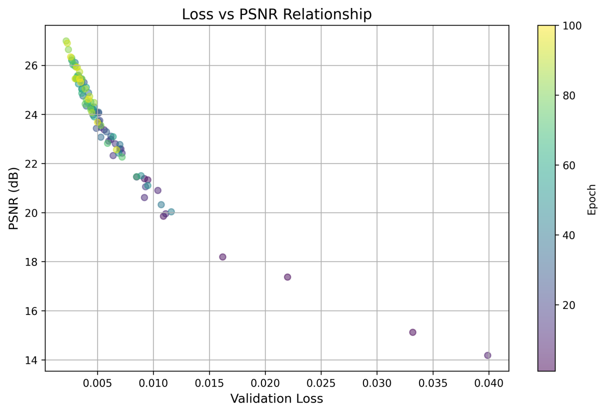

Fig. 4. Loss vs PSNR Relationship

to exploding gradients, gradient clipping is used. Further, the scheduler used is a *Cosine Annealing Warm Restarts* with the parameters $T = 10$,$T_{mul} = 2$, and $\eta_{min} = 1e - 6$.

D. Evaluation Metrics

The proposed model's performance is measured based on primarily three evaluation metrics. These include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) and the validation loss. An in-depth comparison of the proposed model is done in the Experimental Results section of the paper. For visual analysis several plots are also given in the same section.

V. Experimental Results

The enhanced performance of the proposed HybridFUsion-Net is demonstrated using experimentation and comparison with the state-of-the art underwater image enhancement methods such as UGAN, Water-Net, DUIENet, UWCNN, UIEC2-Net, and FUnIE-GAN.

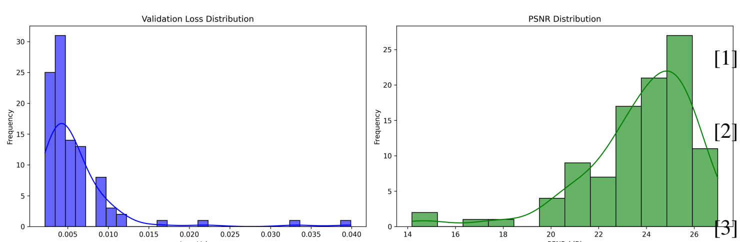

Fig. 5. Validation Loss Distribution and PSNR Distribution

The proposed HybridFusionNet achieves the peak PSNR of 26.99 dB, on UIEB dataset for underwater image restoration. This demonstrates a significant improvement in the quality of reconstructed images in terms of PSNR, SSIM and validation loss. The hybrid fusion model performs at par with other state-of-the and exceeds current state-of-the-art models with PSNRs as DUIENet (24.15 dB), Water-Net (23.82 dB), and UIEC2-Net (23.95 dB). The PSNR of 25.31 dB, also shows the generation high-fidelity restored images even after the first peak performance, demonstrating the robustness of the hybrid approach. The standard deviation of 2.29 dB is also comparatively low and the average PSNR of 23.74 dB during the training procedure shows stability. This suggests the dependable and steady learning during the training .The plots 2, 4, 5 highlight the Hybrid Fusion Network's notable improvements and consistent performance.

At the end of final training, a loss of 0.0570 in achieved in the training phase. The validation loss also got stabilised at a significantly lower 0.0035, indicating the generalisability of the model, free of overfitting. The comparison of performance on fresh, untested data, with other models is given in Table II demonstrating a comprehensive comparison between the models.

In addition to achieving the maximum PSNR among these models,the proposed Hybrid Fusion Model HybridFusionNet has a competitive amount of parameters (2.8M). This shows the model's effective and efficient fusion architecture, that integrates attention-guided multiscale feature fusion, temporal LSTM, and a stacked autoencoder under a perceptual-aware composite loss function.

VI. Conclusion

In this work we propose a hybrid fusion model called HybridFusionNet for Underwater Image Restoration.We utilize the spatial benefits of convolutional neural networks and temporal features using LSTM and use their hybrid for enhanced image restoration. The results demonstrate the superior trade-off between accuracy and model complexity of HybridFusionNet. The performance is evaluated using the PSNR values and validation loss, where the lower validation loss and comparatively high PSNR values demonstrate the well-suitedness of the method for underwater image restoration where both performance and computational cost are critical.

References

[1] Y. Gao, J. Wang, H. Li, and L. Feng, "Underwater image enhancement and restoration based on local fusion," *Journal of Electronic Imaging*, vol. 28, no. 4, pp. 043014–043014, 2019.

[2] X. Guo, X. Chen, S. Wang, and C.-M. Pun, "Underwater image restoration through a prior guided hybrid sense approach and extensive benchmark analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, 2025.

[3] Y. Jin, E. Tan, L. Li, G. Wang, J. Wang, and P. Wang, "Hybrid traffic forecasting model with fusion of multiple spatial toll collection data and remote microwave sensor data," *IEEE Access*, vol. 6, pp. 79211–79221, 2018.

[4] L. Féral, H. Sauvageot, L. Castanet, and J. Lemorton, "Hycell—a new hybrid model of the rain horizontal distribution for propagation studies: 1. modeling of the rain cell," *Radio Science*, vol. 38, no. 3, 2003.

[5] S. K. Hamed, M. J. Ab Aziz, and M. R. Yaakub, "Improving data fusion for fake news detection: A hybrid fusion approach for unimodal and multimodal data," *IEEE Access*, 2024.

[6] J. Wang, R. Wang, X. Zhou, E. Sui, and S. Lu, "A hybrid-structure soft sensor based on model risk assessment and weight fusion for batch processes," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022.

[7] B. L. Ramani, P. Poosapati, P. Tumuluru, C. Saibaba, M. Radha, and K. Prasuna, "Deep learning and fuzzy rule-based hybrid fusion model for data classification," *International Journal of Recent Technology and Engineering*, vol. 8, no. 2, pp. 3205–3213, 2019.

[8] Y. Zang, D. Zhou, C. Wang, R. Nie, and Y. Guo, "Ufa-fuse: A novel deep supervised and hybrid model for multifocus image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–17, 2021.

[9] P. Wen, Y. Li, S. Chen, and S. Zhao, "Remaining useful life prediction of iiot-enabled complex industrial systems with hybrid fusion of multiple information sources," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9045–9058, 2021.

[10] X. Wei, L. Zhang, Y. Lin, J. Wei, C. Zhang, and W. Yi, "Transformer-based multi-sensor hybrid fusion for multi-target tracking," in *2024 27th International Conference on Information Fusion (FUSION)*, pp. 1–8, IEEE, 2024.

[11] D. Li, X. Jia, and J. Zhao, "A novel hybrid fusion algorithm for low-cost gps/ins integrated navigation system during gps outages," *Ieee Access*, vol. 8, pp. 53984–53996, 2020.

[12] S. Yan, X. Chen, Z. Wu, M. Tan, and J. Yu, "Hybrur: A hybrid physical-neural solution for unsupervised underwater image restoration," *IEEE Transactions on Image Processing*, vol. 32, pp. 5004–5016, 2023.

[13] D. L. Muthuraman and S. M. Santhanam, "Visibility improvement of underwater turbid image using hybrid restoration network with weighted filter," *Multidimensional Systems and Signal Processing*, pp. 1–26, 2022.

[14] H.-H. Chang, "Single underwater image restoration based on adaptive transmission fusion," *IEEE Access*, vol. 8, pp. 38650–38662, 2020.

[15] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *2018 IEEE international conference on robotics and automation (ICRA)*, pp. 7159–7165, IEEE, 2018.

[16] M. A. Syariz, C.-H. Lin, M. V. Nguyen, L. M. Jaelani, and A. C. Blanco, "Waternet: A convolutional neural network for chlorophyll-a concentration retrieval," *Remote Sensing*, vol. 12, no. 12, p. 1966, 2020.

[17] H. Zhu and R. Song, "Underwater image enhancement for automotive wading," in *2021 13th International Conference on Computer and Automation Engineering (ICCAE)*, pp. 5–9, IEEE, 2021.

[18] Z. Wang, L. Shen, M. Yu, Y. Lin, and Q. Zhu, "Single underwater image enhancement using an analysis-synthesis network," *arXiv preprint arXiv:2108.09023*, 2021.