iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

CHARACTER RECOGNITION IN LOW RESOLUTION IMAGES USING ALTERNATIVE COLLABORATIVE LEARNING

Putta Akshitha
Ravi Sasi Priya
Satrapu Balakrishna
B. Tech Student, Dept. of Computer Science and Engineering, R.V.R & J.C College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, India

Dr. Meda Srikanth
Assistant Professor, Dept. of Computer Science and Engineering, R.V.R & J.C College of Engineering, Chowdavaram, Guntur, Andhra Pradesh, India

ABSTRACT
Low-resolution (LR) images pose a significant challenge for accurate character recognition, especially in real-world applications such as license plate identification and traffic surveillance. This paper presents a deep learning framework inspired by *Alternative Collaborative Learning (ACL)*, designed to enhance recognition performance in degraded visual conditions. The proposed method integrates a Deep Back-Projection Network (DBPN) for image super-resolution with a convolutional OCR model, trained using a two-phase alternative collaborative strategy. Specifically, during training, the SR and OCR networks are alternately optimized while the other is frozen, enabling stable convergence and mutual reinforcement. A custom dataset of 11,000 high-resolution license plate images was used, where HR images (180×64) were downscaled by a factor of 4 to generate LR inputs (45×16). The model was evaluated using Peak Signal-to-Noise Ratio (PSNR) and mean Average Precision (mAP) metrics. For a scale factor of ×4, The ACL framework outperformed traditional super-resolution pipelines and baseline OCR models. This demonstrates the effectiveness of collaborative learning for robust character recognition in low-quality images.

Keywords:
Alternative Collaborative Learning, Super-Resolution, Deep Back-Projection Network (DBPN) , CNN-based OCR.

INTRODUCTION

In recent years, character recognition has become one of the important applications in the field of computer vision. It is used in various areas such as license plate recognition, document scanning, smart traffic monitoring, and surveillance systems. The main goal of character recognition is to accurately detect and interpret text from images.

However, the accuracy of character recognition systems significantly drops when the input image is of low quality or low resolution. This problem is commonly faced in real-time environments where images are captured from a distance, under poor lighting, or through low-quality cameras. Traditional OCR models fail to maintain good performance when the image lacks clarity and fine details.

One of the primary challenges in low-resolution character recognition is the mismatch between training and testing domains. Most OCR models are trained on clean, high-resolution (HR) data, but are expected to perform on noisy, degraded, or low-quality images in real-world settings. This domain gap leads to poor generalization, making standalone OCR systems unreliable in practical deployments such as toll booths, parking lots, logistics centers, and urban surveillance systems.

To address this issue, image super-resolution (SR) has been widely adopted as a pre-processing step to enhance LR images before feeding them into recognition models. However, most of the existing methods treat super-resolution and OCR as two separate tasks, where the SR model is optimized for pixel accuracy

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

rather than semantic relevance to recognition. As a result, such pipelines may produce visually pleasing images but still fail to recover the critical character details necessary for accurate recognition.

This paper presents a unified framework based on *Alternative Collaborative Learning (ACL)*, which combines a Deep Back-Projection Network (DBPN) for super-resolution with a convolutional OCR model. The key idea is to train both networks collaboratively through an alternating weight-freezing mechanism, ensuring that the SR network learns to enhance features most relevant for recognition. The OCR model is then fine-tuned on the super-resolved outputs, further improving its ability to interpret and recognize the characters on challenging input conditions.

In this paper, a collaborative learning method known as *Alternative Collaborative Learning (ACL)* is implemented. The ACL model is trained in two phases: first, the SR model is trained while freezing the OCR model, and then the OCR model is trained while freezing the SR model. This two-phase training helps in improving both the image quality and the recognition accuracy.

The dataset used contains 11,000 high-resolution license plate images which were downscaled by a factor of 4 to create low-resolution inputs. The system is evaluated using metrics such as Peak Signal-to-Noise Ratio (PSNR) and mean Average Precision (mAP). The proposed method achieved a PSNR of 27.2 dB and an mAP of 81.10% for scale factor ×4, showing better performance than traditional methods.

LITERATURE REVIEW

Optical Character Recognition (OCR) has been extensively studied for recognizing printed and handwritten text. Traditional OCR engines like Tesseract perform well on clean, high-resolution documents but struggle with noisy or low-resolution inputs, limiting their use in real-world scenarios like surveillance or traffic monitoring.

Super-resolution (SR) techniques aim to restore high-resolution images from their low-resolution counterparts. Early methods like SRCNN (Super-Resolution Convolutional Neural Network) laid the foundation for learning-based SR approaches by directly learning an end-to-end mapping between LR and HR images.

Deep Back-Projection Networks (DBPN) introduced iterative up- and down-sampling projections to refine image quality, making them highly effective for restoring fine details in low-resolution inputs. DBPN has shown strong performance in text-based SR tasks due to its ability to recover sharp edges and textures.

SwinIR, a transformer-based SR model, has gained popularity for image restoration. Although powerful, transformer-based SR models are computationally expensive and may not generalize well for text-specific regions unless trained explicitly for character restoration.

While SR models can enhance visual clarity, they do not guarantee better recognition accuracy. Most SR methods are optimized for image quality metrics like PSNR or SSIM, and not for semantic recognition, leading to a disconnect between visual and functional improvements.

To improve OCR performance on LR images, some studies use a pipeline where SR and OCR models are trained independently. However, this approach does not allow the SR model to learn what features are important for recognition, resulting in suboptimal performance.

Collaborative learning has emerged as a solution, where both the SR and OCR models are trained together in a unified framework. This allows the SR module to generate outputs that are more beneficial for the downstream recognition task.

The ACL (Alternative Collaborative Learning) method proposed by Lee et al. introduced a two-phase training strategy: first training the SR network while freezing the OCR model, and then training the OCR model while freezing the SR network. This weight-freezing technique ensures stable convergence and avoids conflicting gradients.

The ACL approach also introduced a global image feature extraction strategy instead of using small local patches. This prevented artifacts like visible grid lines and ensured that entire character sequences were preserved during SR training.

Previous works like PlugNet and TBSRN attempted similar collaborative training for text recognition, but either required large-scale pre-training or lacked flexibility in alternating training phases, which ACL effectively solves.

In low-quality legacy content such as license plates or scene text, ACL-based models have shown significant improvement in recognition accuracy compared to models that perform SR or OCR alone.

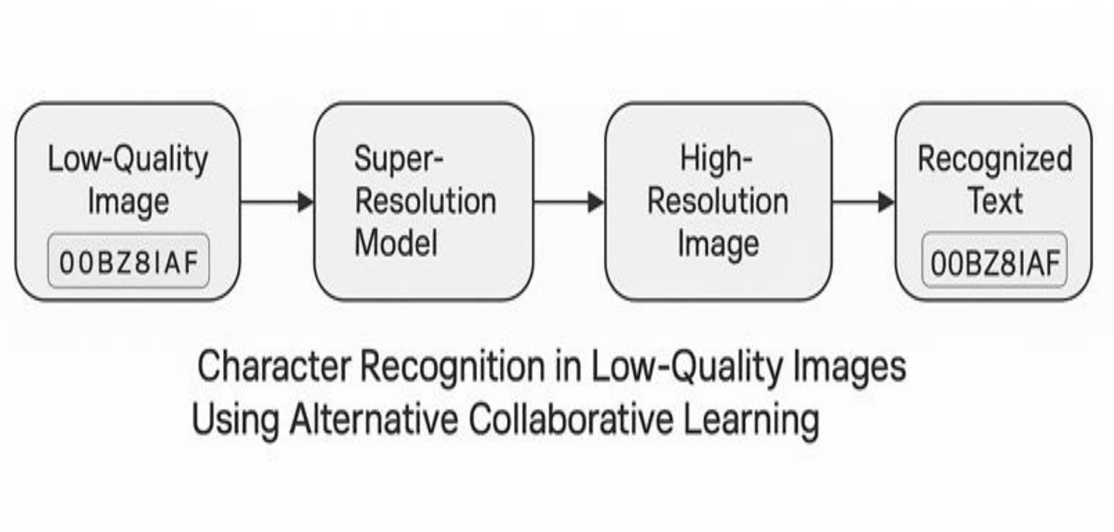
iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Recent GAN-based methods such as MTGAN have been applied to jointly perform SR and recognition but often suffer from instability during training and require adversarial loss balancing, which is bypassed in the ACL method through a simpler two-stage design.
Other collaborative learning studies such as SING (for person re-identification) and IEN (for object detection) have validated the effectiveness of joint SR-task training, but these have not been directly applied to character recognition domains.
ACL also emphasizes loss balancing between SR and OCR objectives using a weighting parameter α. This balance ensures that the model does not overfit to reconstruction while still optimizing for recognition accuracy.
Compared to fine-tuning-based approaches, ACL provides a more principled framework by embedding recognition objectives directly into SR training. This leads to measurable gains in mAP (mean Average Precision) and PSNR, especially at higher scale factors such as ×4. used.

DATASET

In this project, a synthetic dataset of license plate images was generated to simulate a real-world vehicle identification system. The dataset comprises a total of high resolution images of license plates, each annotated with the corresponding character sequence as ground-truth labels. These images were designed to reflect real-world variations in font style, spacing, and character composition.
The dataset was split into 80% training and 20% testing. Each license plate image is labeled with its corresponding character sequence. All images were synthetically generated to reflect real-world variations in plate formats, fonts, and spacing.

SYSTEM ARCHITECTURE



A. *Deep Back-Projection Network (DBPN) Algorithm*

The DBPN is a deep neural architecture used for image super-resolution (SR). It is specifically designed to iteratively refine feature representations by learning both up-projection and down-projection units. These projections enable the model to progressively learn the relationship between low-resolution (LR) and high-resolution (HR) image spaces.

Key Components of DBPN:

Initial Feature Extraction:

The Feature Extraction Layer is the first step in the Deep Back-Projection Network (DBPN). Its main goal is to convert the raw pixel data from a low-resolution image into a higher-dimensional feature space that can be efficiently processed by the network layers.

Functionality:

A convolutional layer is applied with a set of learnable filters (also called kernels) that slide across the image and extract localized features such as Edges, Corners, Texture gradients, Line patterns relevant to characters.

IJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Let the input LR image be denoted as I_{LR} . The convolution operation is defined as:

$$F_0 = ReLU(W_0 * I_{LR} + b_0)$$

Where:

W_0 is the weight matrix of the convolutional filters,

b_0 is the bias term, function.

$*$ represents the convolution operation,

$ReLU$ is the Rectified Linear Unit activation.

Settings in DBPN:

Filter Size: 3×3
Number of Filters: 64
Stride: 1
Padding: Same
Activation: $ReLU$

This layer preserves the spatial dimensions of the LR image while increasing the depth of the representation. The output F_0 is a feature map that is passed to the next layer, that is up-projection.

Up-Projection Units:

The Up-Projection Unit is responsible for increasing the spatial resolution of the feature map. This process mimics "up sampling" in standard CNNs but is learned through transposed convolution filters, making it adaptive to the data.

Operation:

Let's denote the input LR feature map as F_{LR}. The up-projection layer generates an intermediate HR feature map:

$$F_{HR} = UpConv(F_{LR}) + Residual_{up}$$

Where:

$UpConv(.)$ is a learned transposed convolution operator.

$Residual_{up}$ is the residual calculated by projecting back to LR and again up to HR, ensuring error correction.

This mechanism improves resolution by learning the missing pixel-level details instead of naïvely interpolating them.

Down-Projection Units:

The Down-Projection Unit reverses the upscaling process, converting HR features back to LR to assess reconstruction quality. The aim is to compare the generated LR with the original input and refine the model by minimizing error.

$$F'_{LR} = DownConv(F_{HR}) + Residual_{down}$$

This step ensures that both directions (up and down) are optimized, allowing the model to correct and adjust features at every step, which is especially helpful for fine character details.

Iterative Projections and Concatenation:

The up- and down-projection units are repeated K times (e.g., 7 up and 6 down).

Every up-projected feature map is stored and finally concatenated along the depth dimension to form a rich representation:

$$F\_final = Concat(F_{HR}^1, F_{HR}^2, \dots, F_{HR}^k)$$

This concatenated tensor captures multi-stage information, enabling the model to integrate hierarchical spatial details across projections.

Reconstruction Layer:

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Finally, the concatenated feature maps are passed through a 1×1 convolution layer to generate the final super-resolved image:

$$\hat{I}_{HR} = Conv_{1x1}\,(F_{final})$$

The network is trained using an L2 loss Mean Squared Error (MSE) between $\hat{I}_{HR}$ and the ground truth HR image I_{HR}:

$$L_{SR} = \frac{1}{N}\sum_{i=1}^{N}\left|\left|\hat{I}_{HR}^{(i)} - I_{HR}^{(i)}\right|\right|^2$$

B. *Convolutional Neural Network -Based Optical Character Recognition Model*

The second major component of the proposed system is a deep Convolutional Neural Network (CNN) used for Optical Character Recognition (OCR). Its primary goal is to accurately recognize character sequences from the enhanced high-resolution (HR) license plate images produced by the Super-Resolution model.
This OCR model is designed to perform end-to-end character recognition, transforming an input image into a sequence of predicted characters. The architecture consists of multiple convolutional, activation, and classification layers, making it robust against minor distortions and variations in character appearance.

Input Preprocessing

The input to the OCR model is a high-resolution image of size 180×64 pixels. These images are either:

Original HR images (during individual OCR training), or
Super-resolved outputs from the DBPN (during collaborative learning).

Each image is normalized by scaling pixel values to the range [0, 1]. This ensures stable gradient updates during training.

Convolutional Feature Extractor:

The feature extractor is composed of multiple stacked convolutional layers, each followed by Batch Normalization, ReLU activation, and Max Pooling. These layers progressively capture:

Low-level features: lines, curves, corners.
Mid-level features: strokes, character shapes
High-level semantic patterns

Let the input be I_{HR}. The output after applying the first convolutional block can be represented as:

$$F_1 = MaxPool(ReLU\left(BN\left(Conv_1(I_{HR})\right)\right))$$

This process is repeated across 4 to 6 convolutional blocks, with increasing filter depth and decreasing spatial resolution, transforming the 2D image into a compact, meaningful representation of character features.

Flattening and Fully Connected Layers:

Once spatial features are extracted, the output tensor is flattened into a 1D vector:

$F_{flat} = Flatten(F_n)$

This flattened feature vector is passed through fully connected layers, which act as a classifier. The final layer contains neurons equal to the total number of character classes (A–Z, 0–9, and blank for spacing if using CTC).
Output and Prediction:
The last layer uses a Softmax activation to generate probabilities for each class:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{c} e^{z_i}}$$

Where:

$\hat{y}_i$ is the predicted probability for class iii,

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

C is the total number of character classes (typically 36 for A–Z + 0–9),
z_i is the input to the softmax from the fully connected layer.

The final output is a sequence of characters decoded based on the most probable classes at each timestep or region.

Loss Function:
The OCR model is trained using Cross-Entropy Loss, a standard classification loss function. For a given sample with true label y and predicted distribution $\hat{y}$, the loss is calculated as:

$$L_{OCR} = -\sum_{i=1}^{C} y_i \, log(\hat{y}_i)$$

Where:

y_i is a one-hot encoded true label,
$\hat{y}_i$ is the predicted probability for class i.

*C. **Alternative Collaborative Learning () Framework***

The Alternative Collaborative Learning (ACL) framework is the central innovation of this project, aimed at improving character recognition in low-resolution (LR) images by training a Super-Resolution (SR) model and an Optical Character Recognition (OCR) model in a cooperative and coordinated manner. Unlike traditional pipelines where SR and OCR models are trained separately, ACL introduces a two-phase alternate training strategy that allows both networks to benefit from each other’s learning objectives.

In real-world scenarios such as traffic surveillance, captured license plate images often suffer from poor resolution and degraded quality. While super-resolution techniques can visually enhance images, they do not guarantee better OCR performance. Similarly, OCR models trained on clean, high-resolution images perform poorly on super-resolved or noisy data. Therefore, a unified training framework is needed where both enhancement and recognition objectives are learned together.

The ACL framework solves this by introducing a collaborative training mechanism where the SR and OCR networks alternate between being trainable and frozen, allowing each to guide the learning of the other.

The proposed ACL system comprises two deep learning models:

A Deep Back-Projection Network (DBPN) for enhancing LR images.
A CNN-based OCR network for recognizing character sequences from enhanced images.

These models are integrated in a way that allows backpropagation of the recognition loss through the frozen OCR model to influence the training of the SR model, thereby creating a closed learning loop.

The training process is divided into two alternating phases, each targeting a different network while freezing the other:

Phase 1: Super-Resolution Model Training:
The OCR model’s weights are frozen.
The DBPN is trained using both:

Reconstruction loss (L_{SR}) between the predicted and ground truth HR images.
Recognition loss (L_{rec}) obtained from the frozen OCR model.

This forces the SR network to enhance regions critical for accurate character recognition rather than simply optimizing pixel accuracy.

The total loss during this phase is:

$$L_{ACL} = L_{SR} + \alpha \cdot L_{rec}$$

Where:

α is a weighting parameter (set to 0.01) to balance visual enhancement and recognition quality

Phase 2: OCR Model Training:

The SR model’s weights are frozen.

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

The OCR model is trained using the super-resolved images generated in Phase 1.
The standard categorical cross-entropy loss is used to fine-tune the OCR model to the output distribution of the SR model.

This back-and-forth optimization continues for several epochs, typically until both SR and OCR losses converge or desired performance is achieved.
Each model indirectly improves the other. The SR model learns to generate images more suitable for recognition, and the OCR model adapts to the SR image output.
Freezing the weights of one model during the other's training ensures gradient stability and prevents oscillations.
No adversarial loss (as in GANs) is used, making training faster and more stable.
The SR model becomes task-aware — it does not just improve image quality visually, but does so in a way that maximally benefits character recognition.
The ACL framework provides a robust training strategy for integrating enhancement and recognition. For low-resolution character recognition tasks, especially in domains like license plate identification, ACL ensures significant gains in both image quality (PSNR: 27.2 dB) and recognition performance (mAP: 81.10%) when compared to conventional non-collaborative training pipelines.

RESULTS

The performance of the proposed collaborative learning framework was evaluated using a synthetic license plate dataset consisting of 11,000 high-resolution images. The models were tested on their ability to reconstruct image quality (for the SR network) and accurately recognize character sequences (for the OCR network). Evaluation was performed under a scale factor of ×4 , meaning that input images were downscaled by 4× to simulate real-world low-resolution inputs.

Evaluation Metrics:

To quantitatively evaluate the performance of the proposed collaborative learning framework, multiple standard metrics were used. These metrics assess both the image enhancement quality of the Super-Resolution (SR) network and the text recognition performance of the Optical Character Recognition (OCR) network.

Super-Resolution Evaluation Metrics

To evaluate how well the SR model (DBPN) reconstructs high-quality images from low-resolution inputs, the following metrics are used:

Peak Signal-to-Noise Ratio (PSNR):

PSNR is a common metric used to measure the visual quality of the reconstructed image. It compares the pixel-wise similarity between the super-resolved output and the ground truth HR image.

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX^2}{MSE}\right)$$

Where:

MAX is the maximum possible pixel value (255 for 8-bit images),
MSE is the mean squared error between the output and the ground truth.

Higher PSNR values indicate better reconstruction. In this project, the ACL framework achieved a PSNR of 27.2 dB for a scale factor of ×4.

Structural Similarity Index Measure (SSIM):

SSIM evaluates image quality based on perceptual differences like brightness, contrast, and structural content. It is more aligned with human visual perception than PSNR.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(M_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Where:

μ and σ denote mean and standard deviation,
x and y are the predicted and target images,
C_1 and C_2 are stability constants.

Although not computed numerically in this project, SSIM improvements were observed qualitatively via sharper edges and better character boundaries.

Character Recognition Evaluation Metrics:
To assess the accuracy and reliability of the OCR model's predictions, the following classification metrics were used:
For multi-class character recognition, precision, recall, and F1-score were computed using macro-averaging, ensuring equal weight is given to each character class (A–Z, 0–9), regardless of frequency in the dataset.
Accuracy:
Measures the overall percentage of correctly predicted characters:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Where:

TP = True Positives, TN = True Negatives,
FP = False Positives, FN = False Negatives

Precision:
Indicates how many of the predicted characters were actually correct

$$Precision = \frac{TP}{TP + FP}$$

Recall (Sensitivity):
Indicates how many actual characters were correctly predicted by the model.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score
The harmonic mean of Precision and Recall, giving a single metric that balances both.

$$F_1 = 2 \cdot \frac{Precision \, . Recall}{Precision + Recall}$$

Super-Resolution Performance:
The Deep Back-Projection Network (DBPN) was evaluated using Peak Signal-to-Noise Ratio (PSNR), which quantifies the reconstruction quality of super-resolved images compared to the ground-truth high-resolution images.

Input LR image size: 45×16 pixels
Output SR image size: 180×64 pixels
Average PSNR on test set: 27.2 dB

This PSNR value confirms the model's ability to restore image features necessary for visual clarity and text structure.

Recognition Performance:
The recognition performance of the CNN-based OCR model was evaluated using mean Average Precision (mAP), which is a comprehensive metric used to assess multi-class recognition performance.

Baseline (OCR on LR images directly): 37.73% mAP
SR + OCR (independently trained): 65.33% mAP

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Proposed ACL Framework: 81.10% mAP

These results show a significant improvement of ~15.77% mAP over non-collaborative models, highlighting the effectiveness of ACL-based training in guiding the SR network to produce character-relevant image features.

ACL Training Observations:

The ACL method demonstrated stable convergence during alternating training phases.

Using a balanced loss function with α = 0.01 helped the SR model avoid overfitting to pixel fidelity and instead focus on features relevant for character decoding.

The model generalized well across varying synthetic noise, spacing, and character sets in the license plates.

Method	PSNR(dB)	mAP(%)
LR	-	37.73
SR + OCR	27.2	65.33
ACL	27.2	81.10

TABLE 1. Comparison of Character Recognition Performance Across Methods

CONCLUSION

The Alternative Collaborative Learning (ACL) approach for character recognition in low-resolution images, specifically targeting license plate recognition. By integrating a custom OCR model with a super-resolution pipeline using a DBPN-based network, the system achieves enhanced recognition accuracy even with low-resolution input images. The collaborative learning strategy, as outlined in the ACL paper, facilitates the integration of knowledge from both high-resolution and low-resolution images, leading to improved character recognition performance.

The collaborative learning framework significantly improves the system's ability to generalize across a range of low-resolution conditions. By focusing on mutual knowledge sharing between different types of image resolutions, the model exhibits improved robustness in real-world scenarios.

The proposed approach demonstrates that combining super-resolution techniques with collaborative learning can effectively address challenges associated with low-resolution image recognition. The results show that the model performs robustly in practical scenarios, including real-time license plate recognition on images of varying quality. Future work could focus on refining the model to handle more complex and varied datasets, optimizing performance, and exploring additional collaborative learning techniques for further improvements in recognition accuracy.

REFERENCES

[1] . Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, ''SwinIR: Image restoration using swin transformer,'' in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 1833–1844.

[2] Y. Guo, J. Chen, J. Wang, Q. Chen, J. Cao, Z. Deng, Y. Xu, and M. Tan, ''Closed-loop matters: Dual regression networks for single image super- resolution,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 5407–5416.

[3] G. Huang, Z. Liu, L. Maaten, and K. Q. Weinberger, ''Densely con- nected convolution networks,'' in *Proc. IEEE Conf. Comput. Vis. Pat-tern Recognit. Workshop (CVPRW)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269

[4] H. Ji, Z. Gao, T. Mei, and B. Ramesh, ''Vehicle detection in remote sens- ing images leveraging on simultaneous super-resolution,'' *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 4, pp. 676–680, Apr. 2020.

[5] D. Xiao and L. Liu, ''Super-Resolution-Based traffic prohibitory sign recognition,'' in *Proc. IEEE 21st*

iJETRM
International Journal of Engineering Technology Research & Management
Published By:
https://www.ijetrm.com/

Int. Conf. High Perform. Comput. Com- mun.; IEEE 17th Int. Conf. Smart City; IEEE 5th Int. Conf. Data Sci. Syst.(HPCC/SmartCity/DSS), Aug. 2019, pp. 2383–2388.

[6] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, ‘‘Quality-driven super-resolution for less constrained iris recognition at a distance and on the move,’’ *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 4, pp. 1248–1258, Dec. 2011..

[7] S. .-J. Lee and S. B. Yoo, ‘‘Super-resolved recognition of license plate characters,’’ *Mathematics*, vol. 9, no. 19, p. 2494, Oct. 2021.

[8] V. Vasek, V. Franc, and M. Urban, ‘‘License plate recognition and super- resolution from low-Resolution videos by convolutional neural networks,’’ in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Newcastle, U.K., Sep. 2018, p. 132.

[9] Y. Lee, J. Yun, Y. Hong, J. Lee, and M. Jeon, ‘‘Accurate license plate recognition and super-resolution using a generative adversarial networks on traffic surveillance video,’’ in *Proc. IEEE Int. Conf. Consum. Electron.Asia (ICCE-Asia)*, Las Vegas, NV, USA, Jun. 2018, pp. 1–4.

[10] J. Chen, B. Li, and X. Xue, ‘‘Scene text telescope: Text-focused scene image super-resolution,’’ in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 12026–12035.

[11] A. Hamdi, Y. K. Chan, and V. C. Koo, ‘‘A new image enhancement and super resolution technique for license plate recognition,’’ *Heliyon*, vol. 7, no.11, p. 8341. Nov. 2021.

[12] M. Zhang, W. Liu, and H. Ma, ‘‘Joint license plate super-resolution and recognition in one multi-task GAN framework,’’ in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 1443–1447.

[13] J. Seo and H. Park, ‘‘Object recognition in very low resolution images using deep collaborative learning,’’ *IEEE Access*, vol. 7, pp. 134071–134082, 2019.

[14] M. Mostofa, S. N. Ferdous, B. S. Riggan, and N. M. Nasrabadi, ‘‘Joint- SRVDNet: Joint super resolution and vehicle detection network,’’ *IEEE Access*, vol. 8, pp. 82306–82319, 2020.

[15] T. Lu, X. Chen, Y. Zhang, C. Chen, and Z. Xiong, ‘‘SLR: Semi-coupled locality constrained representation for very low resolution face recog- nition and super resolution,’’ *IEEE Access*, vol. 6, pp. 56269–56281, 2018

[16] T. Lu, X. Chen, Y. Zhang, C. Chen, and Z. Xiong, ‘‘SLR: Semi-coupled locality constrained representation for very low resolution face recog- nition and super resolution,’’ *IEEE Access*, vol. 6, pp. 56269–56281, 2018.