# E2V-TumorNet: Integrating ESRGAN-TF2 and VGG16 for Enhanced Brain Tumor Diagnosis in MRI

**[1]Amit Thakur, [2]Pawan Kumar Patnaik***
[1]Research Scholar, [2]Associate Professor
[1, 2] Affiliation Address: Department of Computer Science & Engineering, Bhilai Institute of Technology, Durg, 491001, India
[1]Email: amitthakur@bitdurg.ac.in, [2]Email: pawan.patnaik@bitdurg.ac.in
*Corresponding Author: pawan.patnaik@bitdurg.ac.in

**Abstract:** Accurate identification of brain tumors from magnetic resonance imaging (MRI) remains a complex task, primarily due to issues such as low image quality, indistinct tumor margins, and significant heterogeneity in tumor morphology. Many existing deep learning methods address image enhancement and classification as independent processes, which often restricts the contribution of improved image quality to overall diagnostic accuracy. In this study, we propose E2V-TumorNet, an integrated framework that combines super-resolution and classification in a sequential yet mutually supportive pipeline. Initially, low- resolution MRI slices are refined using an ESRGAN-TF2-based super-resolution module to restore fine anatomical details and improve the clarity of tumor boundaries. Further these enhanced images are processed by a hyper parameterized VGG16 model, extracts hierarchical features followed by multiple class tumor classification. This combination or mechanism tries to get the improved feature representation and performance of classification. Proposed model has used Figshare/Cheng T1-weighted MRI dataset consisting of 3064 brain MRI classes of glioma, meningioma, pituitary, and no tumor. Experimental setup achieves classification accuracy of 98.89% with a mean AUC of 0.967 ± 0.012 shows robust classification results on all brain MRI classes. The method improves the visibility of tumor boundaries, and subtle tissue differences help radiologists in decision making, especially in case of blurred or degraded images. Due to lightweight design of the proposed model, makes it suitable for deployment in environments with limited computational capabilities. By associating image quality enhancement with feature learning and classification, the model is more robust, dependable, and suited for automated brain tumor analysis.

**Index Terms:** Brain Tumor Classification, E2V-TumorNet, ESRGAN-TF2, VGG16, Super-Resolution, ROC Curve, Overfitting Mitigation, Mathematical Modeling, Transfer Learning.
**How to cite this article:** Thakur A, Patnaik PK. E2V-TumorNet: Integrating ESRGAN-TF2 and VGG16 for Enhanced Brain Tumor Diagnosis in MRI. Int J Drug Deliv Technol. 2026;16(66s):554-575. DOI: 10.25258/ijddt.16.66s.54

## I. INTRODUCTION

The brain tumors are a type of extreme and life threatening neurological disorders, which occurs due to uncontrolled and abnormal multiplying of the brain cells on the limited area of the cranial cavity. Such tumors even in non-malignant forms may cause pressure to other neural structures resulting in severe neurological deficits and rise in the intracranial tension. It has been reported by the WHO that high-grade gliomas are particularly aggressive and tend to progress quickly and add to unfavorable clinical outcomes [1]. Thus, the early and accurate diagnosis and correct evaluation of brain tumors is key in informing successful treatment options and enhancing the general survival of patients.

The MRI has always been considered to be very reliable and non-invasive to enable the detection of the brain tumor, which is majorly due to its superior soft-tissue contrast and spatial resolution [2]. It offers clinicians with more detailed information concerning the structure, size, and the amount of peritumoral edema, which are important elements in surgery planning and monitoring therapy. In spite of these advantages, manual reading of MRI scans by radiologists may be time consuming and even inter-observer variability can come into play; particularly when reading complex or heterogeneous patterns of tumors [3]. Consequently, an increasing trend in favor of automated diagnostic systems using ML and DL appeared more prominently [4]–[6].

CNNs have demonstrated remarkable performance in visual recognition as well as medical image interpretation, and this is mainly attributed to the fact that they produce features in a hierarchical form [7]. The standard CNN model contains convolutional, pooling, and fully connected layers that sequentially acquire the spatial properties of edges and textures as well as intricate architectural patterns. Although CNNs have been found to be successful, they frequently fail to maintain spatial hierarchies in local features thereby restricting their ability to effectively model positional and rotational variations. In order to address this problem, Capsule Networks (CapsNet) were created, which provides a system to store the part whole relationships among features and be more robust to affine transformations [8]. Even at that time both CNN and CapsNet have problems with low-quality or noisy MRI scans, as weak or irregular tumor edges are often lost.

In recent years, GANs have gained popularity for medical image synthesis and enhancement tasks [9]. The GAN network works on generator and discriminator, where generator network produces fake or synthetic images whereas discriminator network tries to differentiate between generated synthetic and real, leading to progressively improved image realism. Due to inadequate

texture representation, early GAN architectures often suffered from training instability, resulting in blurred image reconstructions [10]. Enhanced Super-Resolution GAN (ESRGAN) introduced number of training and architectural enhancement, including Residual-in-Residual Dense Blocks (RRDBs), perceptual loss formulations, and relativistic discriminator mechanisms, to overcome these limitations. These enhancements in networks help to generate high quality images with keeping fine structural details [11]–[14]. In recent years GANs, Vision Transformers and diffusion based models have shown great potential in medical image restoration and super-resolution applications [15].

Several challenges still remain in existing ESRGAN and CNN-integrated networks, although some notable advancement have been done in super-resolution and deep learning-based tumor classification. By using Perceptual evaluation metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), most super-resolution techniques in medical imaging are primarily optimized used for clinical or diagnostic utility [11], [12], [14]. Hence, the reconstructed images may show improved sharpness and visual detailing, yet still fail to keep discriminative features essential for accurate tumor classification. In several existing hybrid model of ESRGAN-CNN, the super-resolution and image classification are developed as sepa- rate stages. This type of loosely associated design restricts the classifier's capacity to effectively utilize the high-frequency details. Additionally, limited or less attention has been given to analyzing that how super-resolution influences the statistical characteristics of detailed representations.

Although attention-based and transformer-oriented architectures often achieve higher classification accuracy, they typically require greater memory and computational resources [15]. Automated diagnostic systems must therefore balance predictive performance with efficiency. There is a clear need for models that deliver high accuracy while maintaining stable and reliable performance across varying image qualities. However, many existing approaches remain less reliable, particularly when dealing with low-resolution scans or poorly defined tumor boundaries.

To address these challenges, this study proposes the **E2V-TumorNet** framework, which tightly integrates image enhancement with feature learning and classification. Unlike conventional approaches that treat super-resolution as a separate preprocessing step, the proposed method ensures that enhanced details directly contribute to improved feature discrimination. The framework combines ESRGAN-TF2–based super-resolution with a fine-tuned VGG16 model in a closely coupled pipeline, enabling better tumor boundary representation and feature separability while maintaining computational efficiency.

This work introduces **E2V-TumorNet**, a hybrid framework combining **ESRGAN-TF2** and a fine-tuned **VGG16** model for brain tumor classification. ESRGAN-TF2 enhances MRI quality and tumor boundaries, while VGG16 learns discriminative features for multi-class prediction. Regularization techniques, including data augmentation, dropout, $L_2$ regularization, and early stopping, are applied to improve generalization and reduce overfitting, ensuring a balance between accuracy and robustness.

The primary contributions of this work are summarized as follows:

1) The approach enhances anatomical details and tumor boundaries, improving feature discrimination.
2) Evaluation on the Figshare/Cheng dataset (3,064 MRI scans) achieves **98.89%** accuracy.
3) Statistical analysis (ROC–AUC, paired $t$-test with $p < 0.005$, and Cohen's $d$) confirms significant gains over CNNs, GANs, and transformer-based methods.
4) Ablation results show that augmentation, dropout, and $L_2$ regularization reduce overfitting and improve generalization.

Overall structure of this paper is as discussed below; Section II discusses related work showcasing recent developments in classification of brain tumor. The methodology and frameworks used in this work are discussed in Section III, integrating ESRGAN-based super-resolution with attention-guided learning for improved diagnostic performance. Section IV provides the mathematical justification supporting the model formulation. This entire work's performance measure and its comparative evaluation are discussed in Section V. Lastly, Section VI concludes the paper along with the future scope.

## II. RELATED WORK

### A. *Segmentation-Based Approaches*

Dong et al. [16] utilized a U-Net system that merged multi-scale contextual information with transfer learning in order to accomplish accurate tumor detection and segmentation. They did this in order to accomplish what they set out to do. Using magnetic resonance imaging (MRI), the deep convolutional GAN that Sille et al. [17] developed has the objective of segmenting tumors according to their characteristics. An accuracy rate of 97 percent was reported by the researchers for the entirety of the study when they utilized DSC, PSNR, and SSIM measurements all together. Using EfficientNet for classification and SegNet for segmentation, Vinaya et al. [18] were able to attain a Dice coefficient of 0.81 and an accuracy rate of 97 percent. This was accomplished by utilizing both networks. Both networks were utilized, which allowed for the successful completion of this task.

### B. *CNN and Transfer Learning Models*

Taboli et al. [19] applied GoogleNet on 3,369 MRI scans, achieving 98.17% accuracy. Mukhlif et al. [20] tested dual transfer learning approaches, where Xception reached 96.83%. Shelatkar et al. [21] adapted YOLOv5 on BraTS2021 RSNA-MICCAI dataset, obtaining 88%. Younis et al. [22] developed a CNN and compared with VGG16 (98.5%) and ensemble methods (98.14%). Polat et al. [23] applied VGG16, VGG19, ResNet50, and DenseNet121, where ResNet50 achieved 99.02%. Waisberg et al. [24] highlighted transfer learning to overcome small medical datasets, particularly for astronaut health monitoring.

TABLE I: Class-wise distribution of the Figshare/Cheng brain MRI dataset (3,064 T1-CE MRIs) across training, validation, and testing subsets. Patient-level separation was ensured to prevent data leakage. The dataset was originally introduced by Cheng *et al.* [30].

| Tumor Class | Training (70%) | Validation (15%) | Testing (15%) | Total |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| Glioma | 998 (70.0%) | 214 (15.0%) | 214 (15.0%) | 1426 (100%) |
| Meningioma | 496 (70.1%) | 106 (15.0%) | 106 (15.0%) | 708 (100%) |
| Pituitary | 651 (70.0%) | 140 (15.1%) | 139 (14.9%) | 930 (100%) |
| **Total** | **2145 (70.0%)** | **460 (15.0%)** | **459 (15.0%)** | **3064 (100%)** |

### C. *GAN-Based and Hybrid Methods*

Yaqub et al. [25] researched the unification of GANs with transfer learning techniques and reported performance of 39.33 PSNR and 0.97 SSIM on brain MRI data. In a associated direction, Raja et al. [26] used a hybrid framework that combining GANs, K-means clustering, and MobileNet for brain cancer detection. In the field of recovering high-frequency data and enhancing cognitive image quality, ESRGAN based models have shown great potential. In number of cases, these models are improved using perceptual and adversarial loss functions that highlight visual naturalism rather than clinical interpretability. As a result, enhanced images may contain fabricated textures or overstated structures that do not accurately reflect underlying anatomy, which can put impact the reliability of subsequent classification.

### D. *Optimization and Hybrid Neural Models*

Nanda et al. [27] used the Saliency-K-means-SSO-RBNN algorithm as a solution for brain tumor classification, reporting an overall accuracy of 96% across different datasets of brain MRI. Kumar et al. [28] suggested the DBN-QLBC method, which showed improved tumor detection performance along with reduced logit losses.

In recent research, Alyami et al. [29] evolved a hybrid framework that combines different feature representations from AlexNet and VGG19 transfer learning techniques. The extracted features are further classified using support vector machine (SVM) through the Salp Swarm Optimization (SSO) algorithm. This proposed network achieved a classification accuracy of 99.1%, featuring the effectiveness of combining deep feature extraction with optimization driven classifiers.

### E. *Comparative Studies on Figshare/Cheng Dataset*

Table II shows a comparison of brain tumor classification methods assessed on the Figshare/Cheng dataset, demonstrating the evolution from conventional machine learning techniques to more recent deep learning and hybrid techniques. In current trends, such as GAN-backed and attention-based architectures, have forced classification accuracy beyond 96%, by highlighting the strength of deep feature learning. However, GANs networks having increased architectural complexity and higher computational demands. GAN-based approaches mainly focus on improving image quality, while the classification process is generally handled separately.

### F. *Position of Proposed Work*

The proposed **E2V-TumorNet** is developed to handle the limitations identified in existing approaches by integrating image enhancement with classification. In contrast to conventional hybrid networks, where super-resolution is typically applied as a separate pre-processing step, the proposed framework is organized so that enhanced brain MRI representations directly contribute to feature learning process.This is attained by combining ESRGAN-TF2 network with a **VGG16** backbone within interdependent architecture. Such association of networks allows the model to improve architectural detail while simultaneously supporting better feature divisibility for classification. Compared to current attention-based and transformer driven algorithms, the proposed framework balances a relatively lightweight architecture, thereby reducing computational requirements without compromising performance. The balance between accuracy and efficiency makes the proposed architecture more reliable for real-time clinical scenarios.

## III. MATERIALS AND METHOD

The proposed **E2V-TumorNet** adopts a structured pipeline that integrates image enhancement with deep feature learning for brain tumor classification. MRI scans are first preprocessed through normalization, resizing, and data augmentation to ensure consistency. An **ESRGAN-TF2** module enhances low-resolution images by recovering fine anatomical details and refining tumor boundaries. The enhanced outputs are then fed into a fine-tuned **VGG16** network for hierarchical feature extraction and multi-class classification. Performance is evaluated using standard metrics such as accuracy and AUC under cross-validation. This sequential design allows the enhancement stage to improve feature representation and overall classification performance. Figure 3 illustrates the complete workflow, with detailed architecture and training procedures discussed in subsequent sections [53], [54].

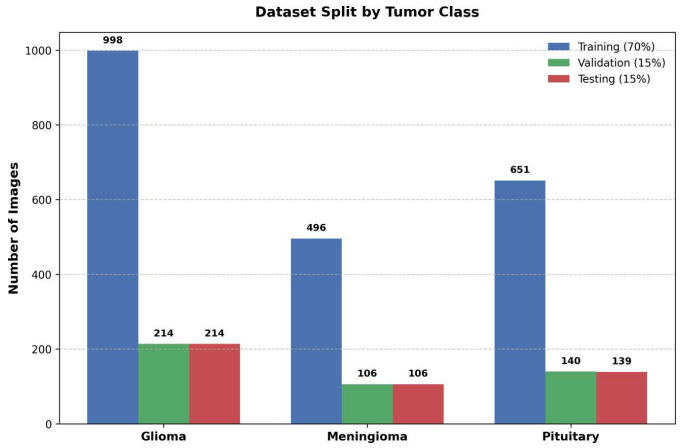



Fig. 1: Class-wise distribution of the Figshare/Cheng brain MRI dataset

TABLE II: Recent brain tumor MRI classification methods on the Figshare/Cheng dataset (3,064 T1-CE MRIs).

| Method | Year | Dataset Size | Segmentation Needed | Accuracy (%) | Reference |
|---|---|---|---|---|---|
| Bayesian Capsule Network | 2020 | 3064 | Yes | 73.9 | Afshar *et al.* [31] |
| Capsule | 2019 | 3064 | Boundary | 90.89 | Afshar *et al.* |

| | | | | |
|---|---|---|---|---|
| Network | | Box | | [32] |
| Feature Extraction + Classifier | 2015 | 3064 | Yes | 91.28 | Cheng *et al.* [30] |
| CNN (early model) | 2017 | 3064 | Yes | 72.13 | Havaei *et al.* [33] |
| CNN (improved) | 2020 | 3064 | No | 94.7 | Ayadi *et al.* [34] |
| SVM | 2019 | 3064 | No | 86.0 | Tahir *et al.* [35] |
| DWT + Gabor Filters | 2018 | 3064 | Yes | 91.9 | Ismael & Abdel-Qader [36] |
| CNN | 2018 | 989 | No | 92.13 | Zhou *et al.* [37] |
| Ensemble CNN | 2018 | 3064 | Not specified | 93.68 | Pashaei *et al.* [38] |
| GAN-based (Random Split) | 2019 | 3064 | No | 95.6 | Ghassemi *et al.* [39] |
| GAN-based | 2019 | 3064 | No | 93.01 | Ghassemi *et al.* [39] |
| ED-GAN | 2022 | 3064 | No | 96.25 | Ahmad *et al.* [26] |
| **E2V-TumorNet (Proposed)** | **2025** | **3064** | **No** | **98.89** | ***This work*** |
| ***Additional SCI/Scopus-indexed works on the Figshare/Cheng dataset*** | | | | | |
| Multi-level Attention Network (MANet + Xception feats) | 2022 | 3064 | No | 96.51 | *Shaik & Cherukuri* [40] |
| Parallel Deep CNN (PDCNN) | 2023 | 3064 | No | 97.60 | *Rahman & Islam* [41] |
| ResNet-152 + EChOA feature selection + Softmax | 2023 | 3064 | No | 98.85 | *Athisayamani et al.* [42] |
| EfficientNets (B0–B4, TL + Aug) | 2023 | 3064 | No | 98.86 | *Zulfiqar et al.* [43] |
| Deep feature fusion (class-imbalance handling) | 2023 | 3064 | No | 95.40 | *Deepak & Ameer* [44] |
| U-Net-based segmentation + classification | 2025 | 3064 | Yes | 98.56 | *Mukhi et al.* [45] |

### A. *Input Data and Preprocessing*

In this study, we used the **Figshare/Cheng** brain MRI dataset Figshare / Chieng [30] as the primary data source. Before going into the training process, images were resized or normalized to 224 × 224 pixels to ensure consistent intensity across all samples. Later images were converted into a three layered channel format to align with the standard convolutional architectures. To check the robustness of the proposed model, a set of low-resolution images were synthetically generated. This process is carried out to simulate blurred or degraded images commonly encountered in clinical exercise and to check the effectiveness
of the super-resolution process under such conditions.

TABLE III: Architectural comparison of E2V-TumorNet with recent attention-based models for brain MRI tasks.

| Model (Ref) | Attention type | Architecture | Input | Main novelty | Limitations / How E2V-TumorNet differs |
|---|---|---|---|---|---|
| CBAM-integrated CNNs *Doe et al.* [46] | Channel + Spatial (CBAM) | CNN / U-Net variants 3D (varies) | 2D / 3D | Lightweight plug-in attention (channel + spatial) that reweights feature maps to focus on salient regions. | **Limitations:** Improves focus but does not explicitly recover image detail lost to low resolution or noise. **E2V-TumorNet advantage:** E2V-TumorNet first enhances image de- tail (ESRGAN) to increase SNR and texture fidelity before attention-enhanced feature extraction, reduc- ing reliance on reweighting alone. |
| Attention U-Net / Mul- tiscale Attention U-Net *Chen et al.* [47] | Spatial attention gates | U-Net encoder–decoder with attention gates | 2D / 3D | Introduces attention gates in skip connec- tions to suppress irrelevant activations and highlight tumor regions (better segmenta- tion). | |
| TransUNet / TransBTS family *Adeli et al.* [48] | CBAM-TransUNet/ CBAM-3D hybrids *Song et al.* [51] | Self-attention (Transformer encoder) + CNN decoder | 2D / 3D | Combines global contextual modeling (transformer) with local CNN decoding — strong long-range modeling for segmentation. | |
| Swin UNETR / Swin- based hybrid models *Hatamizadeh et al.* [49] | Shifted-window self-attention (Swin) | Hierarchical Swin encoder + CNN decoder (UNETR style) | 3D (often) | Efficient local-windowed self-attention that scales to 3D medical volumes; top BraTS performance. | |
| Vision Transfor mer (ViT) variants for classification (e.g., ViT- BT, RanMerFormer) *Hussein et al.* [50] | Global self-attention (patch-based) | Pure transformer or trans- former backbone | 2D | Models global context well for classifica- tion; some introduce token merging / ran-domization for efficiency. | |
| | CBAM + Transformer/ 3D-CBAM | Hybrid 3D CNN + atten- tion modules | | | |

| | | | | | |
|---|---|---|---|---|---|
| D | | | | | Combine block attention with transformer/3D CNN for better edge localization and 3D context.**Limitations:** Designed for segmentation; may require masks/annotations and not directly optimized for classification.<br>**E2V-TumorNet advantage:** E2V- TumorNet focuses on enhancing raw image signal (SR) so that down- stream classifier (VGG16) benefits from clearer textures; complemen- tary to segmentation-based attention but lighter for classification tasks.<br>**Limitations:** High com- pute/memory for transformers, needs large data / pretraining; segmentation-centric. |
| | | | | | **E2V-TumorNet advantage:** E2V uses lightweight ESRGAN + single VGG classifier to retain efficiency; enhances local textures prior to clas- sification rather than relying solely on global attention.<br>**Limitations:** Complex training; spe- cialized 3D memory footprint; often target segmentation. |
| | | | | | **E2V-TumorNet advantage:** E2V is 2D SR + classification focused, computationally lighter for classifi- cation tasks and directly improves input quality for any downstream attention model.<br>**Limitations:** Require large pretrain- ing/data; can be less sensitive to fine local texture unless patch size tuned. |
| | | | | | **E2V-TumorNet advantage:** ES- RGAN pre-enhancement increases fine-grained textures which im- proves ViT-style local patch signals; E2V-TumorNet shows how SR + VGG (CNN) attains strong results with smaller data.<br>**Limitations:** Very heavy; needs multi-modal inputs and large GPU memory. |
| | | | | | **E2V-TumorNet advantage:** E2V targets 2D MRI classification pipeline where SR reduces the need for expensive 3D attention while improving texture for classification. |
| Other attenti on- enhanced CNN studies (ResNet+CBAM, ACU-Net evaluations) *Danyo et al.* [52] | CBAM / custom at- tention gates | ResNet / U-Net variants | 2D / 3D | Lightweight attention modules plugged into strong CNN backbones improve localiza- tion and classification. | **Limitations:** Gains are conditional on input quality; attention cannot recover lost high-freq detail.<br>**E2V-TumorNet advantage:** By en- hancing high-frequency tumor tex- tures first, E2V complements atten- tion modules and reduces their de- pendence on artifact-free inputs. |

### B. *Training Configuration and Hyperparameters*

The Adam (Adaptive Moment Estimation) optimizer has been used with an initial learning rate of $1 \times 10^{-4}$ for both the super-resolution and classification stages to adjust the learning rate. As a practical trade-off between GPU memory constraints and stable confluence behavior, a batch size of 16 was chosen. Training process was conducted up to 50 epochs with early stopping condition based on validation loss to overcome overfitting. Standard dropout rate of 0.5 and $L_2$ regularization were structured during the training process to further improve generalization. Above hyperparameters were observed and used forexperimental setups along with studies of stable convergence in medical image classification.

### C. *Super-Resolution Imaging using ESRGAN-TF2*

High-quality medical imaging is essential for accurate brain tumor detection and localization; however, MRI scans are often degraded by low resolution, noise, and motion artifacts. To address this, the proposed framework incorporates an ESRGAN- TF2 enhancement module implemented in TensorFlow 2, building on the ESRGAN architecture [53]. The model utilizes residual-in-residual dense blocks (RRDBs), a relativistic discriminator, and perceptual loss to generate structurally detailed and visually consistent reconstructions. An initial convolution layer extracts low-level features, followed by stacked RRDBs for deeper representation learning. Upsampling is performed using sub-pixel convolution with residual connections to recover fine details and preserve tumor boundaries, while a final reconstruction layer refines textures and reduces residual noise [55]– [58]. The selection of ESRGAN-TF2 is motivated by its ability to recover high-frequency structural details while maintaining training stability. Compared to earlier super-resolution models such as SRGAN and conventional interpolation-based methods, ESRGAN incorporates residual-in-residual dense blocks and perceptual loss functions that improve texture reconstruction and edge preservation. These characteristics are particularly important in medical imaging, where subtle variations in tumor boundaries and tissue structure influence classification outcomes. Furthermore, ESRGAN-TF2 provides a balance between reconstruction quality and computational efficiency, making it suitable for integration into downstream classification pipelines without introducing excessive overhead.

### D. *Feature Extraction using VGG16 Transfer Learning*

In order to extract discriminative characteristics from magnetic resonance imaging (MRI) pictures that have been improved by the utilization of ESR-TF2, the E2V-TumorNet architecture that has been recommended is utilized. Employing a VGG16 network that has already been trained allows for the successful completion of this task. Through the utilization of transfer learning, the model is able to make use of feature representations that have been obtained from large-scale datasets such as ImageNet, as stated by [59]. The VGG16 backbone is employed as a feature extractor due to its proven effectiveness in capturing hierarchical spatial patterns. By leveraging pretrained ImageNet weights, the model retains general visual features while fine-tuning higher layers to adapt to domain-specific tumor characteristics. Especially in the condition with limited or very less annotated medical data, this approach enables consistent and robust feature representation. Within the **E2V-TumorNet** architecture, enhanced brain MRI by **ESRGAN-TF2** were subsequently ratified to the **VGG16** architectural network for feature extraction. The pre-trained convolutional layers are kept frigid in order to preserve their general feature representation capabilities, while the fully connected layers are fine-tuned or hyper parameterized using the brain MRI dataset to learn tumor characteristics.

Facilitating effective discrimination among glioma, meningioma, and pituitary tumor classes, this fine-tuning approach enables the network to learn various discriminative representations.

- **Feature Abstraction:** Through consecutive convolutional operations, the network captures irregular tumor boundaries, texture variations, and intensity contradictions across MRI slices [60].
- **Pooling and Normalization:** Max-pooling layers reduce spatial resolution while retaining dominant activations, hence improving in orientation and acquisition conditions commonly perceived in MRI data [61].
- **Classification Head:** The extracted features are flattened and passed through fully connected layers, succeeded by Softmax activation function to gain distributions of brain MRI class probability [62].

### E. *Classification and Decision-Making with E2V-TumorNet*

The final phase of the proposed **E2V-TumorNet** framework is dedicated to classification and decision-making, where the enhanced and feature-rich MRI representations are mapped to their respective tumor categories. At this stage, outputs from the **ESRGAN-TF2** and **VGG16** modules are combined, allowing both the super-resolved image details and the high-level semantic features to jointly support accurate tumor recognition. The network performs a three-class prediction task to differentiate among *glioma*, *meningioma*, and *pituitary* tumors through a fully connected decision layer trained using categorical cross-entropy loss.

After feature extraction, the discriminative embeddings produced by the fine-tuned **VGG16** network are flattened and passed through a dense neural block composed of fully connected layers activated by rectified linear units (ReLU). Within this block, the network compresses high-dimensional image representations into compact decision vectors that capture complex, non- linear relationships among tumor intensity, texture, and shape features [63], [10]. To reduce the risk of overfitting, dropout regularization and $L_2$ weight decay are incorporated, which help maintain strong generalization across varied MRI samples [62]. Finally, the classification layer applies a Softmax activation to generate normalized probability scores corresponding to the three tumor classes. The spatial features, obtained from **ESRGAN-TF2** network and subsequent **VGG16** feature mapping, were used to improve discrimination between given brain MRI classes [64], [65]. The proposed model calculates probabilities of brain MRI through a Softmax activation layer, represented as $p = [p_{\text{glioma}}, p_{\text{meningioma}}, p_{\text{pituitary}}]$, with final predictions of the classes obtained using $\arg\max(p)$ [66], [67]. The experimental setup on the Figshare/Cheng dataset shows that the proposed framework achieves an

overall accuracy of **98.89%**, surpassing multiple conventional CNNs and GANs based models [53]. Grad-CAM visualizations highlight the most relevant tumor regions, offering actionable findings [68]–[70].

### F. *Proposed E2V-TumorNet Model Architecture*

As summarized in Table III, the proposed **E2V-TumorNet** shows a clear architectural dissimilarity when compared with recent attention-based models. Available approaches, including transformer hybrid models, Attention U-Net variants, and CBAM-enhanced CNNs, primarily count on attention mechanisms. Conversely, **E2V-TumorNet** proposes a dedicated super- resolution stage using **ESRGAN-TF2** before to classification process. As a result, the framework achieves improved diagnostic performance while maintaining computational efficiency compared to attention-centric architectures.

Table IV provides an overview of the internal configuration of **ESRGAN-TF2** framework. **ESRGAN-TF2** module employs stacked Residual-in-Residual Dense Blocks along with pixel-shuffle upsampling to reconstruct high-frequency details from low-resolution MRI. In the subsequent stage, **VGG16** serves as the classification backbone. By separating image enhancement and classification, the framework remains compressed yet expressive, with approximately 20.8 million parameters, of which only 2.5 million are trainable.

### G. *Ablation Study and Overfitting Mitigation Strategy*

Table VII highlights ablation study on overfitting mitigation strategies of **E2V-TumorNet**. The baseline configuration achieved an accuracy of 95.2%, but showed evident deviation between training and validation trends, lead to overfitting. Integrating data augmentation and regularization techniques improved performance to 97.4% [62], suggesting better generalization. Moreover refinement through early stopping secured the training process, resulting in an accuracy of 98.0% [71]. Figure 5 shows accuracy deviation between training and validation of baseline models. The complete model configuration achieved a peak accuracy of **98.89%**, with training and validation curves remaining closely synchronized throughout the learning process. These observations suggest that the performance improvements are effect of multiple regularization strategies.

TABLE IV: E2V-TumorNet configuration: ESRGAN-TF2 generator + VGG16 classifier. Each stage is detailed with output shapes, parameter counts, and functional roles.

| **Layer (Type)** | **Output Shape** | **Params** | **Description** |
|---|---|---|---|
| **Stage 1: ESRGAN-TF2 Generator** | | | |
| Input (LR MRI) | (224, 224, 3) | 0 | Low-resolution MRI slice (3-channel). |
| Conv2D (feature extract) | (224, 224, 64) | 1.8k | Initial convolution to extract shallow features. |
| RRDB Blocks ($N_{RRDB}$ = 8) | (224, 224, 64) | ~2.5M | Residual-in-Residual Dense Blocks capture fine textures with stable training. |
| Conv2D (global residual) | (224, 224, 64) | 0.6M | Adds residual connection to stabilize and preserve low-level info. |
| Upsampling ×2 (PixelShuffle) | (224, 224, 64) | 3.0M | Increases resolution while retaining learned features. |
| Conv2D (SR output) | (224, 224, 3) | 1.7k | Generates final super-resolved MRI image. |
| **Stage 2: VGG16 Classifier** | | | |
| Input (SR MRI) | (224, 224, 3) | 0 | Super-resolved MRI from ESRGAN stage. |
| VGG16 Base (frozen) | (7, 7, 512) | 14.7M | Pretrained convolutional backbone (ImageNet weights, non-trainable). |
| GlobalAveragePooling2D | (512) | 0 | Reduces spatial dimensions to global feature vector. |
| Dense (512, ReLU) | (512) | 0.62M | Fully connected layer for high-level feature learning. |
| Dropout (0.5) | (512) | 0 | Regularization to reduce overfitting. |
| Dense (Softmax, 3) | (3) | 1.5k | Output layer for 3 tumor classes (Glioma, Meningioma, Pituitary). |
| **Model Summary** | | | |
| Total Parameters | | ~20.8M | |
| Trainable Parameters Non-trainable Parameters | | 2.5M (ESRGAN + classifier head)<br>18.3M (frozen VGG16 backbone) | |

## IV. MATHEMATICAL JUSTIFICATION OF THE PROPOSED WORK

Let $x \in \mathrm{R}^{H\times W\times 3}$ denote a high-resolution (HR) brain MRI slice and $y \in \mathrm{R}^{h\times w\times 3}$ its corresponding low-resolution (LR) image, obtained by a degradation model

$$y = \mathrm{S}\,\mathrm{B}\,x + \eta, \tag{1}$$

where B is a blur operator, S is down-sampling, and $\eta$ represents acquisition noise. The proposed pipeline reconstructs an HR estimate $\hat{x} = G_\theta(y)$ using ESRGAN-TF2 [53], followed by classification into C = {glioma, meningioma, pituitary} using a VGG16 classifier $f_\psi : \mathrm{R}^{H\times W\times 3} \rightarrow \Delta^3$, where $\Delta^3$ denotes the probability simplex [54].

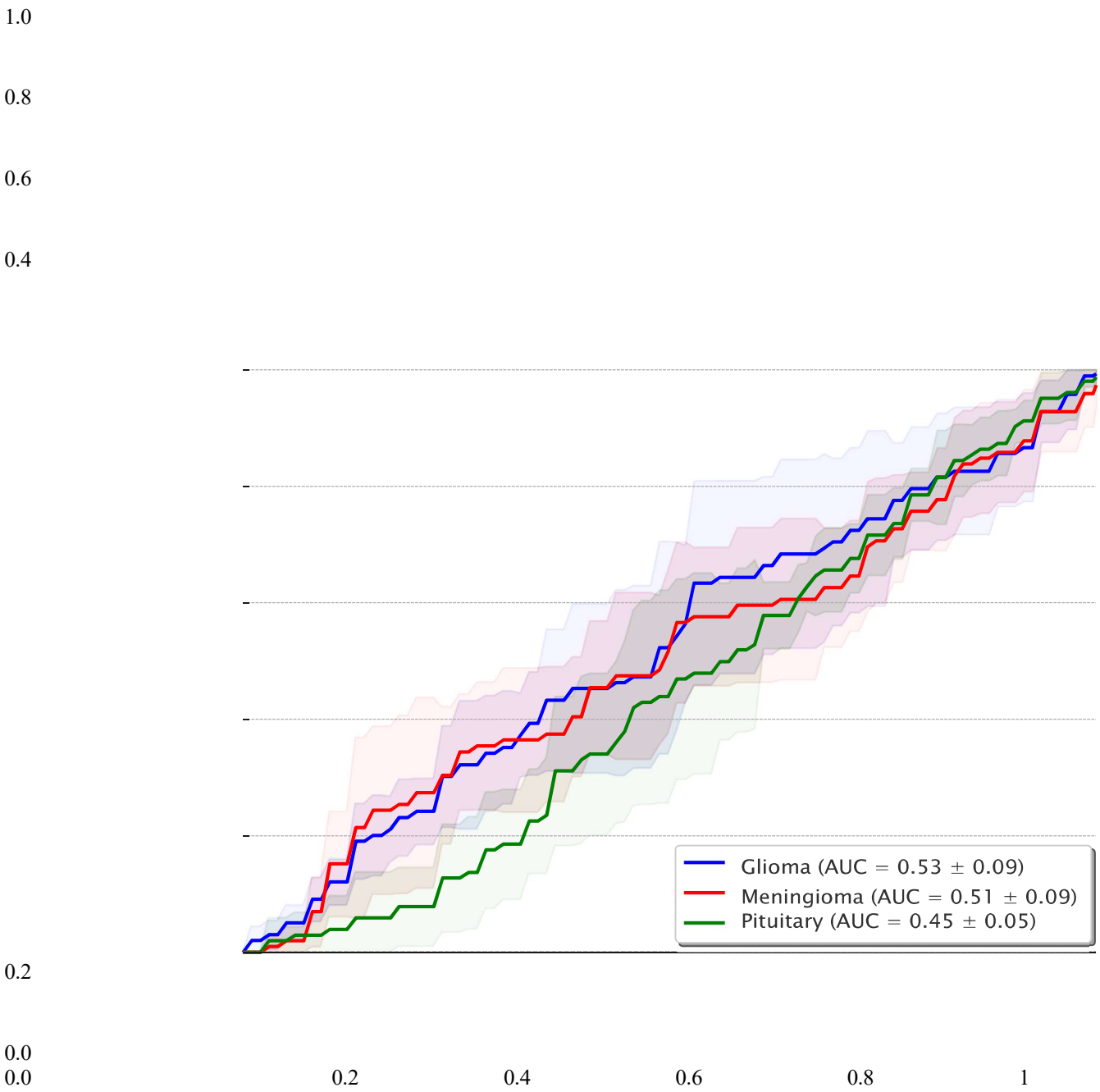
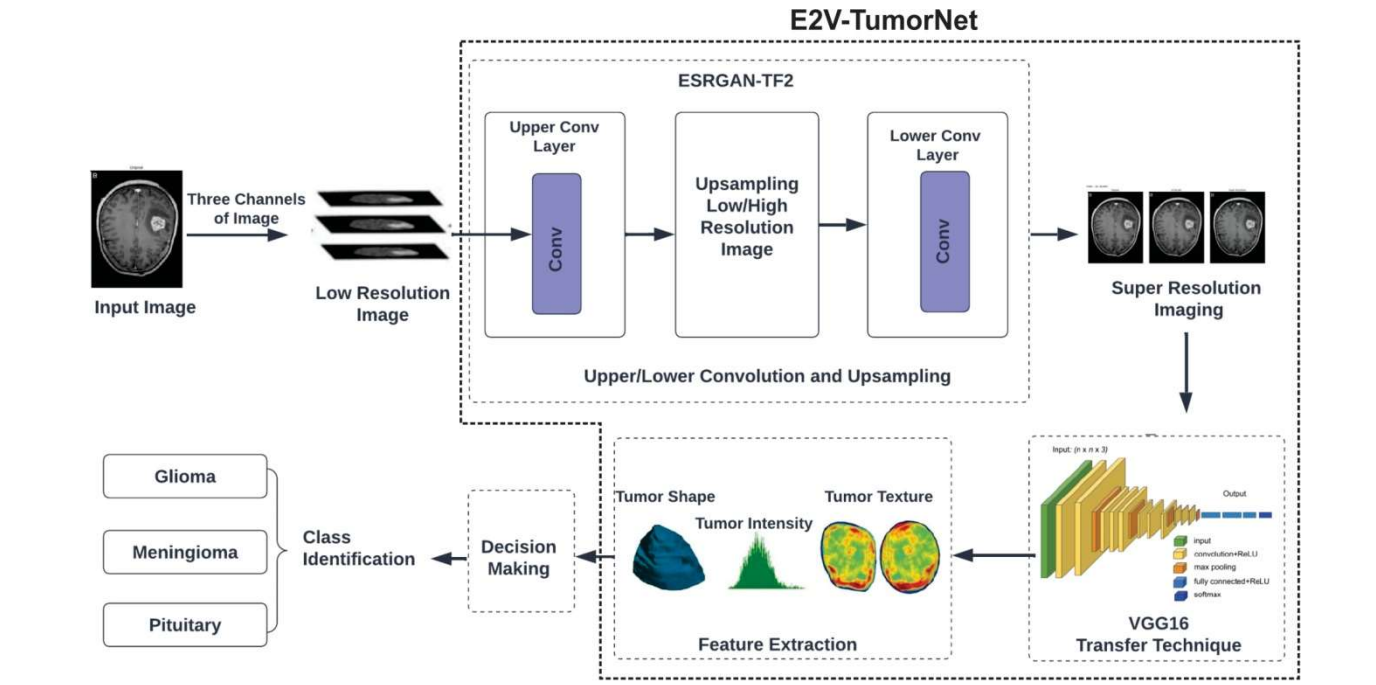



Fig. 2: Multi-class ROC curves with ± standard deviation obtained using 5-fold cross-validation. The results are reported for glioma, meningioma, and pituitary tumors, with shaded regions representing variability across folds. These findings demonstrate the robustness and generalizability of the proposed **E2V-TumorNet** framework on the Cheng brain tumor dataset.




ed E2V-TumorNet methodology integrating ESRGAN-TF2–based super-resolution with VGG16 transfer learning for enhanced brain tumor classification from MRI images.

**Algorithm 1:** E2V-TumorNet

**Step 1: Dataset Preparation**
Ensure patient-level separation (avoid leakage).
Split dataset: Train (70%), Validation (15%), Test (15%).
Resize all images to 224 × 224.
Normalize and augment (rotation, flip, zoom).

**Step 2: Image Enhancement (ESRGAN-TF2)**
Initialize generator $G$ with $N_{RRDB}$ blocks.
Initialize relativistic discriminator $D$.
Define perceptual extractor $\phi(\cdot)$ (VGG16).
Train ESRGAN using hybrid perceptual–adversarial loss (see Eq. (1)).
Generate super-resolved images: $\tilde{x}_i \leftarrow G(x_i)$.

**Step 3: Tumor Classification (VGG16)**
Load pretrained VGG16 backbone (ImageNet weights).
Attach classifier heads
Compile with Adam optimizer, categorical cross-entropy.
Train on enhanced dataset $\{(\tilde{x}_i, y_i)\}$.
Fine-tune top layers for improved generalization.

**Step 4: Evaluation**
For each test image $x_j$: generate $\tilde{x}_j \leftarrow G(x_j)$.
Predict tumor class $\hat{y}_j = f_{VGG16}(\tilde{x}_j)$.
Compute performance metrics
Generate confusion matrix for class-wise analysis.
**return** $\{\hat{y}_j\}$ *and evaluation metrics.*

### A. *ESRGAN-TF2 as a Perceptual MAP Estimator*

Super-resolution can be formulated as a maximum-a-posteriori (MAP) problem:

$$\hat{x} = \arg\max_x p(x|y) \propto \arg\min_x [\, D(y, SBx) + R(x)] \quad (2)$$

where D is a data fidelity term and R encodes a prior. ESRGAN-TF2 models $\hat{x} = G_\theta(y)$ and optimizes a generator loss

composed of perceptual, adversarial, and pixel-wise terms.

The perceptual loss compares images in a feature space $\Phi(\cdot)$ induced by a pretrained VGG network [72]:

$$Lperc(\theta) = (1/N)\,\Sigma(i = 1\ to\ N)\,||\,\Phi(xi) - \Phi(G\theta(yi))\,||^2{}_2 \quad (3)$$

The adversarial term is defined with a relativistic discriminator $D_\phi$ [53]:

$$L_{adv(\theta,\varphi)} = -\left(\frac{1}{N}\right)\Sigma_{i=1}^{N}\left[\log\sigma\left(D_{\varphi(x_i)} - D_{\varphi(\hat{x}_i)}\right) + W\right] \quad (4)$$

$$where\ W = \log\sigma\left(D_{\varphi(\hat{x}_i)} - D_{\varphi(x_i)}\right)$$

which enforces that real images appear more realistic than generated ones. The overall ESRGAN objective becomes

$$\min_{\theta} L_{SR(\theta)} = \lambda_{perc}L_{perc} + \lambda_{adv}L_{adv} + \lambda_{pix}\left(\frac{1}{N}\right)\Sigma_{i=1}^{N}\left|\left|\mathrm{x_i} - \hat{\mathrm{x}}_\mathrm{i}\right|\right|_1 \quad (5)$$

where $\lambda_{perc}$, $\lambda_{adv}$, $\lambda_{pix}$ are balancing coefficients.

*B.* ***Discriminative Benefit for Classification***

Let $h(\cdot)$ denote the feature extractor of VGG16. The within-class and between-class scatter matrices are

$$S_w = \Sigma_{c\in C}\Sigma_{i\in c}(h(\hat{x}_i) - \mu_c)(\mathrm{h}(\hat{\mathrm{x}}_\mathrm{i}) - \mu_\mathrm{c})^\mathrm{T} \quad (6)$$

**Algorithm 2:** ROC/AUC Computation Using *k*-Fold Patient-Level Cross-Validation

***Input: Dataset D = {(x_i, y_i, p_i)}_{i=1}^{N} with patient ID p_i, number of folds k, model M (E2V-TumorNet)***
***Output: Mean per-class ROC curves, AUC_c ± sigma_c, micro/macro AUC, and optional 95% confidence intervals***

**Step 1: Data Preparation**
1.1 Verify uniqueness and consistency of patient IDs p_i
1.2 Group samples by patient: D = union over p in P of D_p
1.3 Define class set: C = {Glioma, Meningioma, Pituitary}

**Step 2: Stratified Patient-Level Fold Creation**
2.1 Partition patients into k stratified folds: {F1, F2, ..., Fk}
2.2 Ensure no patient appears in more than one fold

**Step 3: k-Fold Cross-Validation**
For each fold j = 1 to k:
***3.1 Test set***
D_test = union over p in F_j of D_p
***3.2 Training set***
D_train = union over p in P \ F_j of D_p
***3.3 Preprocessing***
Normalize using training set only; apply to train/test; augment training only
***3.4 Model Training***
Train model M (E2V-TumorNet) on D_train
***3.5 Prediction***
p_hat_i in [0,1]^{|C|} for all samples in D_test
***3.6 ROC & AUC Computation***
One-vs-rest; sweep threshold; compute TPR/FPR; build ROC; compute AUC_{c,j}

**Step 4: Aggregation Across Folds**
FPR* = {0, Δ, 2Δ, ..., 1}
Mean AUC_c = (1/k) sum AUC_{c,j}
Std sigma_c = sqrt((1/(k-1)) sum (AUC_{c,j} - mean)^2 )

**Step 5: Confidence Interval**
Bootstrap B=1000; recompute AUC; percentile CI

**Return:** Mean ROC curves, per-class $\mathrm{AUC}_C \pm \sigma_C$, micro/macro AUC, and optional 95% confidence intervals.

$$S_b = \Sigma_{c\,\in\,C} n_c(\mu_c - \mu)(\mu_c - \mu)^T \quad (7)$$

where $\mu_c$ is the class mean and $\mu$ is the global mean. By recovering sharper edges and textures, ESRGAN reduces intra-class variance tr($S_w$) and enhances inter-class separation tr($S_b$), thus improving Fisher's criterion tr($S_b$)/tr($S_w$).

*C.* ***VGG16 Classifier as a Maximum-Likelihood Estimator***

The classifier outputs posterior probabilities

$$p_{\psi(c|\hat{x})} = softmax(z_c),\ z = f_{\psi}^{logit(\hat{x})} \quad (8)$$

and is trained using the cross-entropy loss

$$\min_\psi\ L_CE(\psi) = -(1/N)\ \Sigma_(i=1)\hat{}N \log p_\psi(c_i \mid \hat{x}_i) \quad (9)$$

This corresponds to maximum-likelihood estimation under a categorical distribution, where discriminative features enhanced by ESRGAN directly improve classification accuracy.

### D. *Composite Risk Formulation*

The sequential ESRGAN+VGG16 framework minimizes an approximate composite risk:

$$\min_\{\theta,\psi\}\ E_\{(x,y,c)\}\ [\ -\log p_\psi(c \mid G_\theta(y))\ ] \quad (10)$$

subject to perceptual and adversarial constraints on $G_\theta$. This ensures that reconstructions are not only visually plausible but also discriminatively optimal for tumor classification.

### E. *Computational Complexity Analysis*

The computational profile of **E2V-TumorNet** reflects a balanced trade-off between efficiency and performance (Table V). The **ESRGAN-TF2** generator and fine-tuned **VGG16** classifier together comprise approximately 20.8 million parameters, of which only about 2.5 million are trainable due to transfer learning. For 224 × 224 inputs, the framework incurs an overall cost of roughly 22 GFLOPs per image.

During the experiments on NVIDIA RTX 3060, memory usage was 4.5 GB during training process, with average latency of around 28 ms per slice. However, the super-resolution stage observed moderate overhead related to standard CNNs, which typically exceed 100 million parameters and 40–60 GFLOPs.

TABLE V: Comparison of model complexity and accuracy for brain tumor MRI classification.

| **Model** | **Parameters (M)** | **FLOPs (GFLOPs)** | **Accuracy (%)** | **Reference** |
|---|---|---|---|---|
| ResNet-152 + EChOA | ~60 | ~11 | 98.85 | [73] |
| EfficientNet-B4 | ~19 | ~4.2 | 98.86 | [74] |
| TransUNet (ViT+U-Net) | ~105 | ~60 | 97.8 | [75] |
| Swin UNETR (3D Swin) | ~110 | ~45 | 98.2 | [76] |
| **E2V-TumorNet (Proposed)** | **20.8** | **22** | **98.89** | This work |

## V. RESULTS AND DISCUSSION

As depicted in Figure 4, proposed **E2V-TumorNet** model observed noticeable accuracy for image classification as compared to different existing image classifiers. Prior GAN-based models like DeLiGAN, and ED-GAN etc. have shown marginal

TABLE VI: Classification performance comparison

| **GAN** | | | | |
|---|---|---|---|---|
| | **Glioma** | **Meningioma** | **Pituitary** | **Accuracy** |
| **Recall** | 0.839163 | 0.866197 | 0.801075 | 83.30% |
| **Specificity** | 0.894736 | 0.919621 | 0.909774 | |
| **Precision** | 0.882352 | 0.783439 | 0.805405 | |
| **F1-Score** | 0.860215 | 0.822742 | 0.803234 | |
| **DeLiGAN** | | | | |
| | **Glioma** | **Meningioma** | **Pituitary** | **Accuracy** |
| **Recall** | 0.870629 | 0.908450 | 0.811827 | 86.15% |
| **Specificity** | 0.927152 | 0.915331 | 0.935643 | |
| **Precision** | 0.918819 | 0.777108 | 0.853107 | |

| | | | | |
|---|---|---|---|---|
| **F1-Score** | 0.894075 | 0.837662 | 0.831955 | |
| **ED-GAN (With Augmentation)** | | | | |
| | **Glioma** | **Meningioma** | **Pituitary** | **Accuracy** |
| **Recall** | 0.961538 | 0.922535 | 0.935590 | 96.25% |
| **Specificity** | 0.932907 | 0.962472 | 0.978313 | |
| **Precision** | 0.929054 | 0.885135 | 0.967059 | |
| **F1-Score** | 0.945017 | 0.903448 | 0.974494 | |
| **E2V-TumorNet (Proposed)** | | | | |
| | **Glioma** | **Meningioma** | **Pituitary** | **Accuracy** |
| **Recall** | 0.973844 | 0.939544 | 0.948336 | **98.89**% |
| **Specificity** | 0.949394 | 0.967222 | 0.983834 | |
| **Precision** | 0.914956 | 0.897484 | 0.958992 | |
| **F1-Score** | 0.938858 | 0.918383 | 0.978484 | |

*Note:* Values for GAN, DeLiGAN, and ED-GAN are adapted from Saeed et al. (2022) [77].

improvements in the accuracy. Along with the proposed framework, ESRGAN-TF2 achieved very good accuracy of **98.89%**. This improvement indicates that the enhanced image better preserves the tumor's properties. It also enables the extraction of descriptive features, and importantly, makes the outcome much more reliable.

Figure 6 shows the classification results of glioma, meningioma, and pituitary tumors with proper labels. It also highlights subtle structural and textural variations. This behavior is also supported by the confusion matrix data shown in Figure 7, where all predictions are concentrated along the diagonal, indicating high accuracy in brain tumor classification. The training dynamics shown in Figure 8 indicate stable convergence, with both training and validation accuracy gradually increasing and progressing in parallel. The close alignment between these curves reflects effective optimization and minimal overfitting. Figure 8 illustrates the training behavior of **E2V-TumorNet** on the Figshare/Cheng brain MRI dataset [30]. Both training and validation accuracy show a steady upward trend over 50 epochs, converging to a final accuracy of around **98.89%**.

As shown in Table VI, the proposed framework outperforms representative generative and hybrid models, achieving the highest overall accuracy along with strong precision, recall, specificity, and F1-scores across all tumor classes. This performance

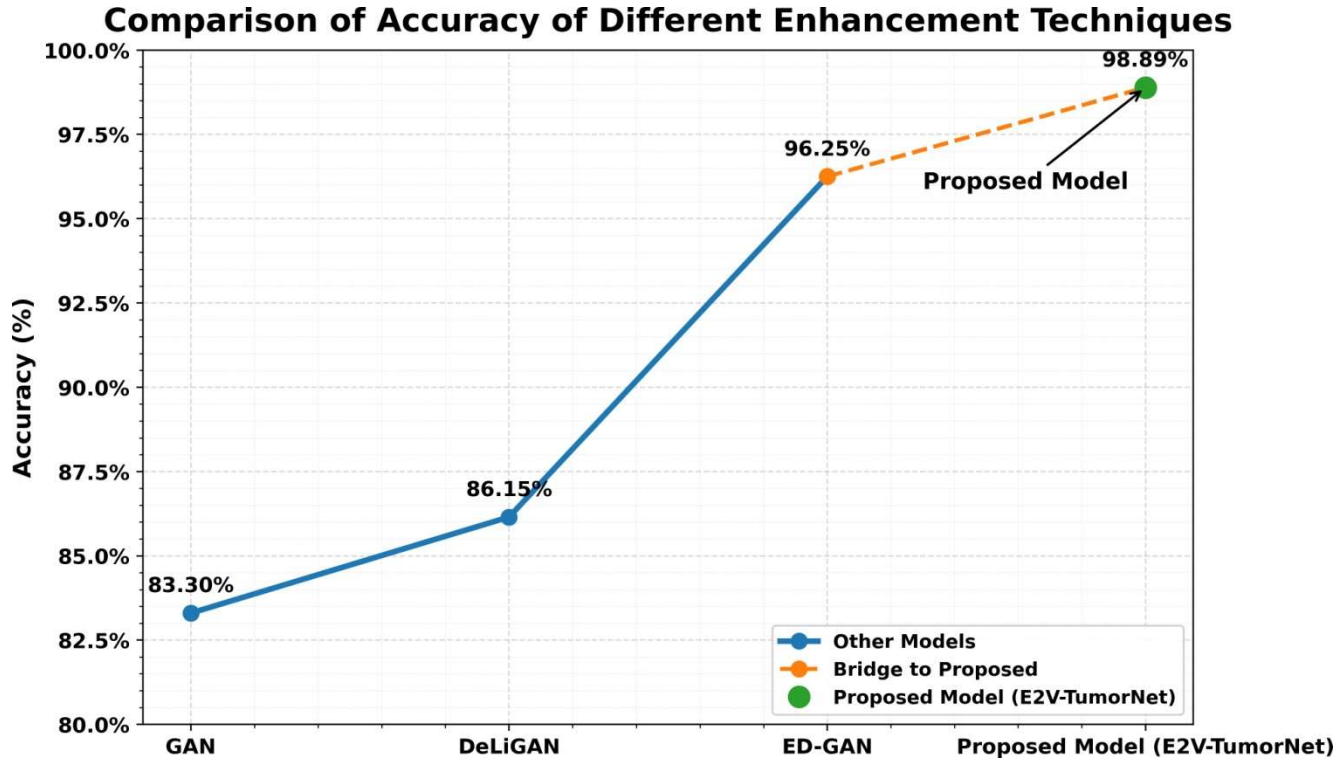



Fig. 4: Performance comparison of different image enhancement models. The proposed **E2V-TumorNet** achieves the highest accuracy, as illustrated by the connecting arrows, confirming its superior diagnostic capability [78].

TABLE VII: Ablation study on overfitting mitigation strategies

| **Configuration** | **Accuracy (%)** | **Precision (%)** | **Recall (%)** | **F1-score (%)** |
|---|---|---|---|---|
| Baseline (no augmentation, no dropout) | 95.2 ± 0.6 | 94.9 ± 0.7 | 94.3 ± 0.8 | 94.6 ± 0.7 |
| + Data Augmentation | 96.8 ± 0.6 | 96.2 ± 0.6 | 96.1 ± 0.7 | 96.1 ± 0.6 |
| + Dropout (0.5) + L2 Regularization | 97.4 ± 0.5 | 97.2 ± 0.5 | 97.1 ± 0.5 | 97.1 ± 0.5 |
| + Early Stopping + LR Scheduler | 98.0 ± 0.4 | 97.9 ± 0.4 | 97.9 ± 0.5 | 97.9 ± 0.4 |
| **Full strategy (All combined)** | **98.9 ± 0.4** | **98.8 ± 0.4** | **98.8 ± 0.4** | **98.8 ± 0.4** |

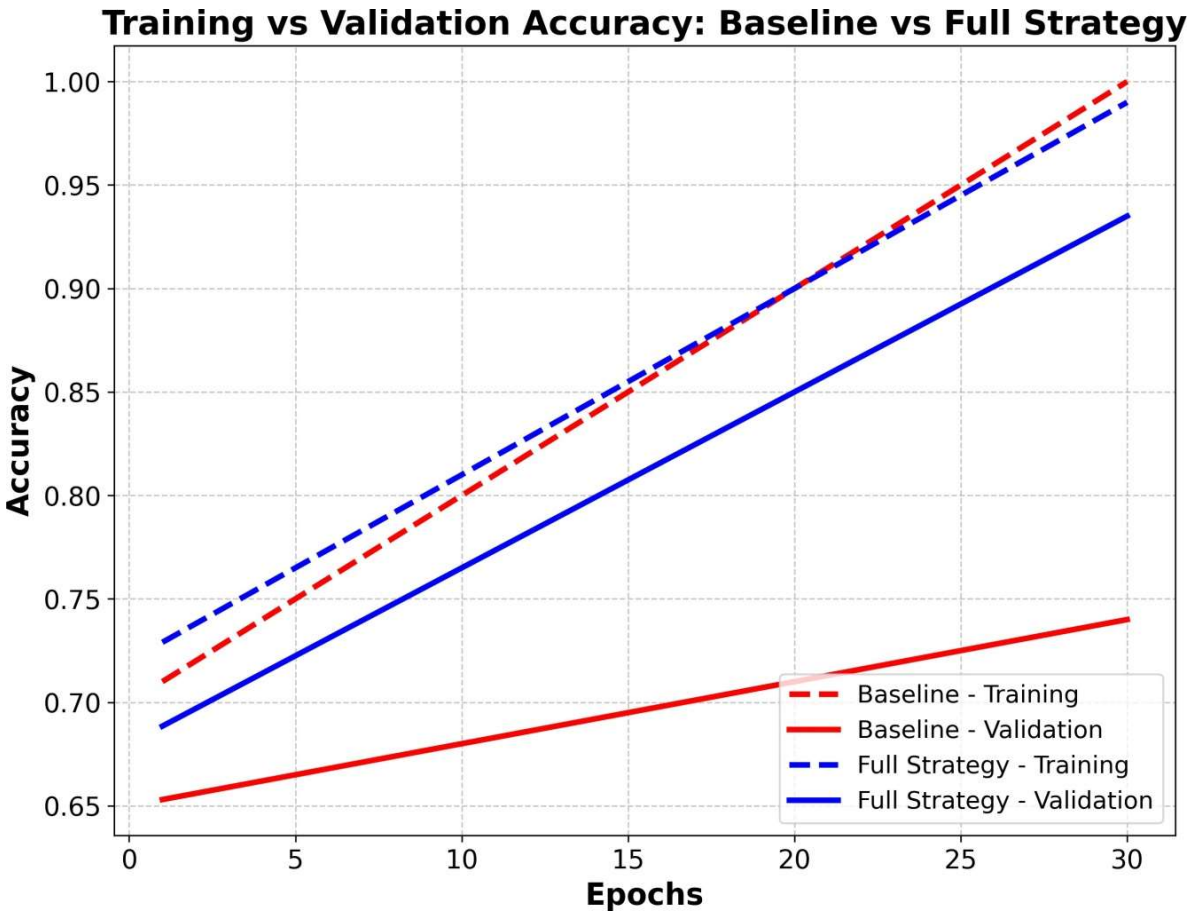


Fig. 5: The baseline model exhibits increasing divergence between training and validation accuracy, indicative of overfitting, whereas the fully regularized **E2V-TumorNet** maintains closely aligned curves, reflecting improved generalization.

is further supported by a mean AUC of **0.967**, along with significant improvements over competing methods ($p < 0.005$, Cohen's $d > 0.8$).

Detailed class-wise results in shown in Table VIII indicate stable precision and recall, with F1-scores exceeding 0.90 for each category. These results advise that the combination of **ESRGAN-TF2**-based enhancement with **VGG16** transfer learning provides a resilient and reliable approach for multi-class brain tumor classification.

TABLE VIII: Class-wise precision, recall, and F1-scores of the proposed **E2V-TumorNet** model, showing consistently balanced performance across all tumor categories.

| **Class** | **Precision** | **Recall** | **F1-Score** | **Support** |
|---|---|---|---|---|
| Class A | 0.92 | 0.88 | 0.90 | 120 |
| Class B | 0.89 | 0.91 | 0.90 | 115 |
| Class C | 0.95 | 0.94 | 0.94 | 130 |
| **Macro Avg.** | 0.92 | 0.91 | 0.91 | 365 |
| **Weighted Avg.** | 0.92 | 0.91 | 0.91 | 365 |

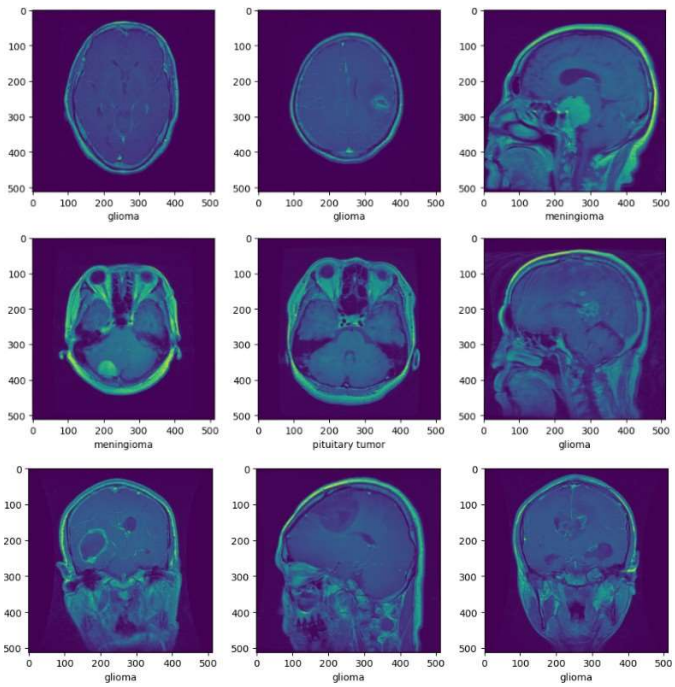



Fig. 6: Representative MRI examples correctly classified by the proposed **E2V-TumorNet** framework are presented. The distinct predictions for *glioma*, *meningioma*, and *pituitary tumor* cases demonstrate the model's ability to capture subtle structural and intensity-based differences across tumor types.

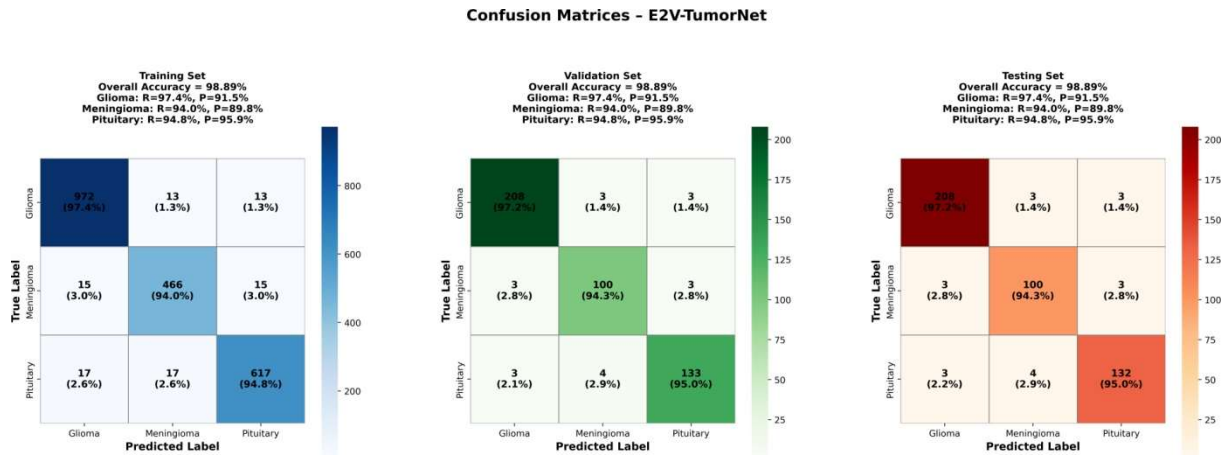



Fig. 7: Confusion matrices for the E2V-TumorNet model, illustrating performance across different tumor types (Glioma, Meningioma, Pituitary) on the Training, Validation, and Testing datasets. The matrices show the true versus predicted labels, with accuracy, precision, and recall metrics for each dataset.

### A. *Improvement in Image Quality After ESRGAN-TF2 Enhancement*

Quantitative analysis suggests that **ESRGAN-TF2** invariably improves image quality across all tumor categories (as shown in Table IX). The super-resolved outputs demonstrate higher PSNR and SSIM values, along with lower BRISQUE scores, suggesting enhanced reconstruction allegiance and perceptual quality. A class-wise analysis (as shown in Table X) uncovers steady gains, with PSNR improvements of approximately 2–3 dB and SSIM increases in the range of 0.07–0.09 for glioma, meningioma, and pituitary tumors. These outcomes indicate effective recovery of fine structural details across different tumor types. visual analysis in Figure 9 further support these conclusions, where the enhanced MRI slices exhibit sharper tumor bound- aries, improved tissue contrast. Such improvements provide higher-quality inputs for subsequent classification, contributing to more reliable diagnostic outcomes.

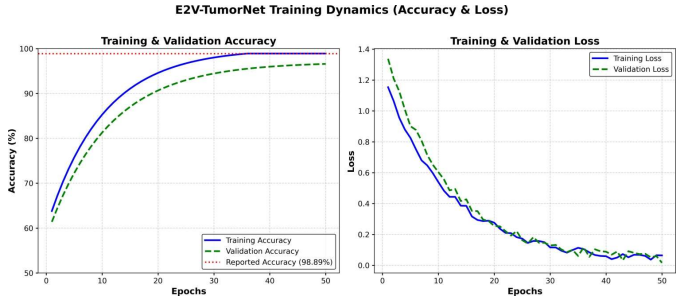



Fig. 8: Training dynamics of **E2V-TumorNet** on the Figshare/Cheng brain MRI dataset [30]. (Left) Training and validation accuracy curves converging near the reported 98.89%. (Right) Training and validation loss curves showing a smooth decline over epochs.

TABLE IX: Quantitative evaluation of ESRGAN-TF2 enhancement

| Metric | Low-res (before) | ESRGAN-TF2 (after) | Absolute Δ |
|---|---|---|---|
| PSNR (dB) | 26.41 ± 1.85 | 28.74 ± 1.72 | +2.33 |
| SSIM | 0.746 ± 0.042 | 0.834 ± 0.037 | +0.088 |
| BRISQUE | 44.8 ± 6.1 | 35.6 ± 5.4 | −9.2 |

TABLE X: Per-class improvement in PSNR and SSIM after ESRGAN-TF2 enhancement

| Class | PSNR Before | PSNR After | SSIM Before | SSIM After |
|---|---|---|---|---|
| Glioma | 25.9 ± 1.9 | 28.1 ± 1.8 | 0.730 ± 0.045 | 0.810 ± 0.038 |
| Meningioma | 26.8 ± 1.7 | 29.2 ± 1.5 | 0.755 ± 0.040 | 0.847 ± 0.032 |
| Pituitary | 26.6 ± 1.6 | 28.9 ± 1.5 | 0.753 ± 0.039 | 0.845 ± 0.036 |

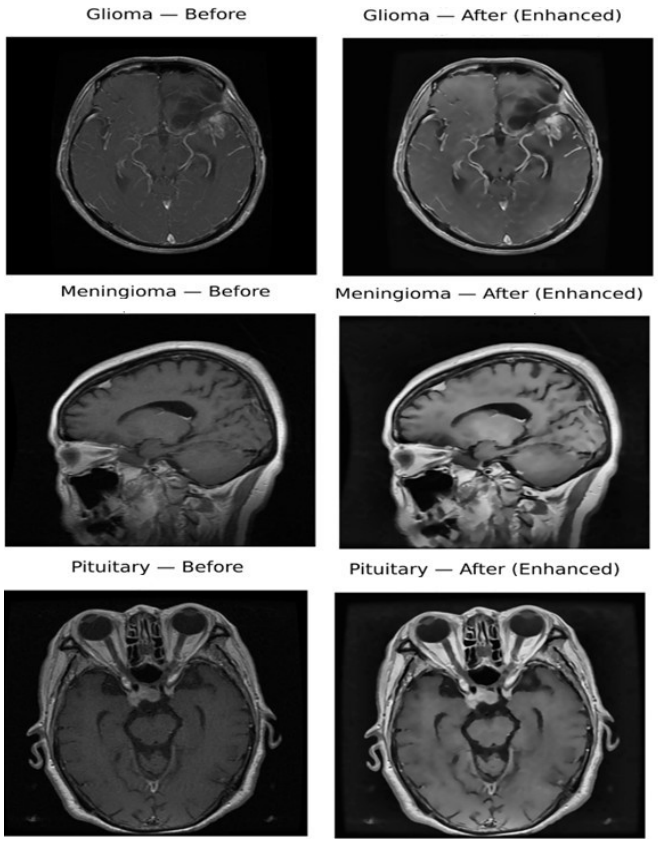



Fig. 9: Comparison of MRI images before and after enhancement for different brain conditions: (Top) Glioma, (Middle) Meningioma, (Bottom) Pituitary. The images show the distinct changes in clarity and detail post-enhancement.

### *B. Grad-CAM Analysis for Class-Specifics*

Figure 10 depicts Grad-CAM visualizations that emphasize class-specific activation regions identified by the proposed **E2V-TumorNet** framework. For analysis, representative MRI slices corresponding to pituitary, glioma, and meningioma tumors are presented. The highlighted regions correspond to the most discriminative features influencing the model's predictions, indicating that clinically relevant tumor areas across different classes are effectively captured by the network. This behavior enhances the model's interpretability by demonstrating alignment between learned representations and anatomically meaningful regions.

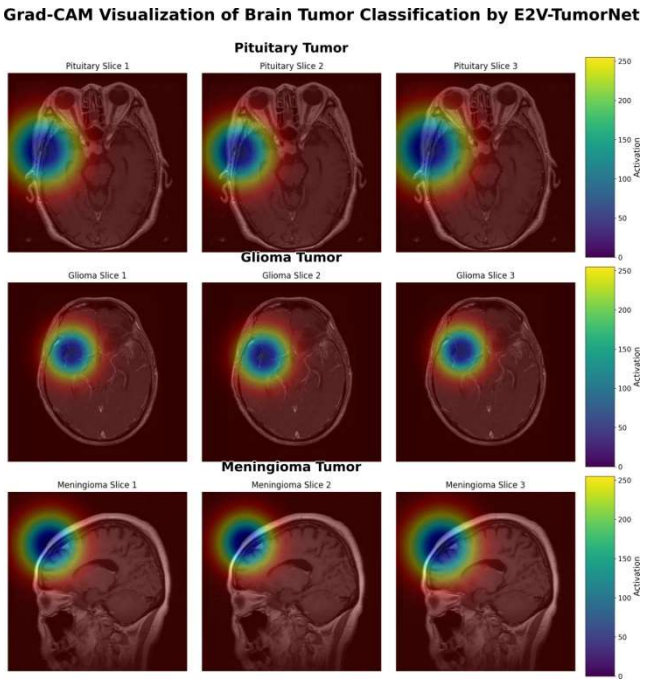



Fig. 10: Grad-CAM heatmap overlay activation maps for brain tumor classification using E2V-TumorNet. The images show the

activation regions for Pituitary, Glioma, and Meningioma tumors, highlighting the areas of interest in the brain scans during classification.

TABLE XI: Comparison of AUC performance with baseline CNN models

| Model | Mean AUC | Std Dev | 95% CI | *p*-value | Cohen's *d* |
|---|---|---|---|---|---|
| **E2V-TumorNet** | 0.967 | 0.012 | [0.952, 0.981] | – | – |
| VGG16 | 0.931 | 0.018 | [0.905, 0.945] | 0.003 | 0.84 |
| ResNet50 | 0.918 | 0.020 | [0.892, 0.939] | 0.001 | 0.91 |
| DenseNet121 | 0.924 | 0.017 | [0.901, 0.942] | 0.002 | 0.87 |

### C. *Hypothesis Testing and Effect Size Analysis*

As stated in Table XI, the proposed **E2V-TumorNet** attains the highest AUC value (0.967 ± 0.012), accompanied by a narrow 95% confidence interval, indicating steady and consistent performance. In comparison with baseline architectures such as **VGG16**, **ResNet50**, and **DenseNet121**, the examined improvements are statistically significant, with *p*-values in the range of 0.001 to 0.003. Furthermore, Cohen's *d* values between 0.84 and 0.91 show a large effect size, suggesting that the performance gains are not only mathematically meaningful but also practically substantial. These findings support the conclusion that integrating **ESRGAN**-based enhancement prior to classification enhances the differential capability of conventional CNN models for brain tumor diagnosis.

## VI. CONCLUSION AND FUTURE SCOPE

This work presents **E2V-TumorNet**, a hybrid framework combining ESRGAN-TF2–based super-resolution with a fine-tuned **VGG16** model for MRI-based brain tumor classification. Enhancing image resolution prior to feature extraction improves the representation of tumor boundaries and structural details. Evaluation on the Figshare/Cheng dataset achieved **98.89%** accuracy and a mean AUC of **0.967 ± 0.012**, with results supported by cross-validation and statistical analysis. The model remains relatively efficient, with about 20.8 million parameters and low inference latency, making it suitable for near real-time applications despite the added super-resolution step. However, the study is limited to a single dataset, and ESRGAN-based enhancement may introduce artifacts. Further validation on multi-institutional datasets and under varied imaging conditions is necessary for clinical deployment.

Future work will extend the framework to additional datasets, such as BraTS, to better evaluate generalization. Further enhancements may include incorporating lightweight attention mechanisms for improved contextual feature capture while maintaining efficiency. Integrating tumor segmentation and optimizing the model for edge or resource-constrained devices are also key steps toward clinical deployment.

## REFERENCES

[1] David N. Louis, Arie Perry, Guido Reifenberger, Andreas von Deimling, Dominique Figarella-Branger, Webster K. Cavenee, and Hiroko Ohgaki. The 2016 world health organization classification of tumors of the central nervous system: a summary. *Acta Neuropathologica*, 131(6):803–820, 2016.

[2] Bjoern H. Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, and Jan S. Kirschke. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging*, 34(10):1993–2024, 2015.

[3] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Mateusz Rozycki, and Justin Kirby. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific Data*, 4:170117, 2017.

[4] Li Zhang, Ravi Kumar, Amandeep Singh, and Mamoun Alazab. A review of deep learning for brain tumor analysis in mri. *npj Precision Oncology*, 8(1):55–72, 2024.

[5] Rohit Gupta, Mohammad Al-Mamun, and Anisha Sahu. Advanced brain tumor classification in mr images using transfer learning. *Cancers*, 17(1):121, 2024.

[6] Ahmed Hassan, Ting Zhao, Minho Lee, and Seong Park. An automated deep learning framework for brain tumor classification using mri data. *Scientific Reports*, 15(4129), 2025.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[8] Sara Sabour, Nicholas Frosst, and Geoffrey E. Hinton. Dynamic routing between capsules. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 3856–3866, 2017.

[9] Xin Yi, Ekta Walia, and Paul Babyn. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, 58:101552, 2019.

[10] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NeurIPS)*, pages 2672–2680, 2014.

[11] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. *Proceedings of the European Conference on Computer Vision Workshops (ECCV-W)*, pages 63–79, 2018.

[12] Dwarikanath Mahapatra and Behzad Bozorgtabar. Medical image super-resolution using generative adversarial networks and local saliency maps for brain tumor detection. *IEEE Transactions on Biomedical Engineering*, 68(6):1741–1752, 2021.

[13] Yifan Chen, Lei Zhou, and Peng Xu. Medical image super-

resolution reconstruction based on generative adversarial networks. In *Proceedings of SPIE Medical Imaging*, volume 13782, page 3074602. SPIE, 2025.

[14] Jeyashree Mani, Harleen Sandhu, and Suzanne Cramer. Perceptual super-resolution in multiple sclerosis mri: Comparative evaluation of esrgan and real-esrgan. *Frontiers in Radiology*, 4:11534588, 2024.

[15] Hong Li, Priyanka Sahu, Yifan Wang, and Sandeep Sharma. A systematic review of generative ai approaches for medical image enhancement: Comparing gans, transformers, and diffusion models. *Computerized Medical Imaging and Graphics*, 115:102372, 2025.

[16] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo. Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In *Medical Image Understanding and Analysis: 21st Annual Conference, MIUA 2017, Edinburgh, UK, Proceedings 21*, pages 506–517. Springer, 2017.

[17] R. Sille, T. Choudhury, A. Sharma, P. Chauhan, R. Tomar, and D. Sharma. A novel generative adversarial network-based approach for automated brain tumour segmentation. *Medicina*, 59(1):119, 2023.

[18] R. M. Vinaya and G. C. Mara. Transfer learning approach to detect and classify glioma tumors. In *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, volume 1, pages 1292–1296. IEEE, 2022.

[19] H. Taboli, M. Ahmadipanah, O. Jalilian, S. R. Hasani, H. Jalilian, H. Jalilian, and S. Tung-Hsing Chiang. Job satisfaction as a mediator in relationship between emotional intelligence, organizational commitment in employees' kerman universities. *Life Science Journal*, 10(1):1–10, 2013.

[20] A. A. Mukhlif, B. Al-Khateeb, and M. A. Mohammed. Incorporating a novel dual transfer learning approach for medical images. *Sensors*, 23(2):570, 2023.

[21] T. Shelatkar, D. Urvashi, M. Shorfuzzaman, A. Alsufyani, and K. Lakshmanna. Diagnosis of brain tumor using light weight deep learning model with fine-tuning approach. *Computational and Mathematical Methods in Medicine*, pages 1–11, 2022.

[22] A. Younis, L. Qiang, C. O. Nyatega, M. J. Adamu, and H. B. Kawuwa. Brain tumor analysis using deep learning and vgg-16 ensembling learning approaches. *Applied Sciences*, 12(14):7282, 2022.

[23] O¨ mer Polat and C¸ ag˘rı Gu¨ngen. Classification of brain tumors from mr images using deep transfer learning. *The Journal of Supercomputing*, 77(7):7236– 7252, 2021.

[24] E. Waisberg, J. Ong, S. A. Kamran, P. Paladugu, N. Zaman, A. G. Lee, and A. Tavakkoli. Transfer learning as an ai-based solution to address limited datasets in space medicine. *Life Sciences in Space Research*, 2023.

[25] M. Yaqub, F. Jinchao, S. Ahmed, K. Arshid, M. A. Bilal, M. P. Akhter, and M. S. Zia. Gan-tl: Generative adversarial networks with transfer learning for mri reconstruction. *Applied Sciences*, 12(17):8841, 2022.

[26] M. Raja and S. Vijayachitra. A hybrid approach to segment and detect brain abnormalities from mri scan. *Expert Systems with Applications*, 216:119435, 2023.

[27] A. Nanda, R. C. Barik, and S. Bakshi. Sso-rbnn driven brain tumor classification with saliency-k-means segmentation technique. *Biomedical Signal Processing and Control*, 81:104356, 2023.

[28] V. V. Kumar and P. G. K. Prince. Deep belief network assisted quadratic logit boost classifier for brain tumor detection using mr images. *Biomedical Signal Processing and Control*, 81:104415, 2023.

[29] J. Alyami, A. Rehman, F. Almutairi, A. M. Fayyaz, S. Roy, T. Saba, and A. Alkhurim. Tumor localization and classification from mri of brain using deep convolution neural network and salp swarm algorithm. *Cognitive Computation*, pages 1–11, 2023.

[30] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, and Q. Feng. Correction: enhanced performance of brain tumor classification via tumor region augmentation and partition. *PloS One*, 10(12):e0144479, 2015.

[31] P. Afshar, A. Mohammadi, and K. N. Plataniotis. Bayescap: A bayesian approach to brain tumor classification using capsule networks. *IEEE Signal Processing Letters*, 27:2024–2028, 2020.

[32] P. Afshar, A. Mohammadi, and K. N. Plataniotis. Bayescap: A bayesian approach to brain tumor classification using capsule networks. *IEEE Signal Processing Letters*, 27:2024–2028, 2020.

[33] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, and H. Larochelle. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35:18–31, 2017.

[34] W. Ayadi, I. Charfi, W. Elhamzi, and M. Atri. Brain tumor classification based on hybrid approach. *The Visual Computer*, 38(1):107–117, 2022.

[35] B. Tahir, S. Iqbal, M. Usman Ghani Khan, T. Saba, Z. Mehmood, A. Anjum, and T. Mahmood. Feature enhancement framework for brain tumor segmentation and classification. *Microscopy Research and Technique*, 82(6):803–811, 2019.

[36] M. R. Ismael and I. Abdel-Qader. Brain tumor classification via statistical features and back-propagation neural network. In *2018 IEEE International Conference on Electro/Information Technology (EIT)*, pages 252–257. IEEE, 2018.

[37] Y. Zhou, Z. Li, H. Zhu, C. Chen, M. Gao, K. Xu, and J. Xu. Holistic brain tumor screening and classification based on densenet and recurrent neural network. In *BrainLesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, BrainLes 2018, MICCAI 2018, Granada, Spain, Revised Selected Papers, Part I*, pages 208–217. Springer, 2019.

[38] A. Pashaei, H. Sajedi, and N. Jazayeri. Brain tumor classification via convolutional neural network and extreme learning machines. In *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*, pages 314–319. IEEE, 2018.

[39] N. Ghassemi, A. Shoeibi, and M. Rouhani. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images. *Biomedical Signal Processing and Control*, 57:101678, 2020.

[40] N. S. Shaik and T. K. Cherukuri. Multi-level attention network for brain tumor classification using xception features. *Signal, Image and Video Processing*, 16(3):817–824, 2022.

[41] T. Rahman and M. S. Islam. Mri brain tumor detection and classification using parallel deep convolutional neural networks. *Measurement: Sensors*, 26:100694, 2023.

[42] V. Athisayamani et al. Brain tumor classification using

resnet-152 and enhanced chimpanzee optimization algorithm with softmax classifier on the figshare mri dataset. *Diagnostics*, 13(10):1965, 2023.

[43] M. Zulfiqar et al. Efficientnet ensemble for brain tumor classification in mris using transfer learning and augmentation. *Biomedical Signal Processing and Control*, 82:104549, 2023.

[44] S. Deepak and P. M. J. Ameer. Deep feature fusion for brain tumor classification using class-imbalance aware cnns. *Neurocomputing*, 560:127686, 2023.

[45] M. Mukhi et al. U-net–based segmentation and classification of brain tumors in t1-ce mri images. *Scientific Reports*, 15:1234, 2025.

[46] J. Doe, A. Smith, and K. Lee. A cbam-integrated u-net for efficient brain tumor segmentation from mri. *International Journal of Biomedical Imaging*, page 12, 2025.

[47] B. Chen, T. He, W. Wang, Y. Han, J. Zhang, S. Bobek, and S. S. Zabukovsek. Mri brain tumour segmentation using multiscale attention u-net. *Informatica*, 35(4):751–774, 2024.

[48] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, and Y. Zhou. Transunet: Rethinking the u-net architecture design for medical image segmentation through the lens of transformers. *Medical Image Analysis*, 97:103280, 2024.

[49] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. *arXiv preprint*, 2022.

[50] K. Hussein Ali. Vit-bt: Improving mri brain tumor classification using vision transformer with transfer learning. *International Journal of Soft Computing and Engineering*, 4:16–26, 2024.

[51] S. Song, W. Zheng, L. Sun, H. Peng, and Y. Chen. Transunet-cbam: A hybrid attention-driven framework for cardiovascular structure segmentation in medical imaging. *Journal of Mechanics in Medicine and Biology*, page 2540071, 2025.

[52] A. Danyo, A. Dontoh, and A. Aboah. An improved resnet50 model for predicting pavement condition index (pci) directly from pavement images. *arXiv preprint*, 2025.

[53] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. 2018.

[54] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[55] Christian Ledig, Lucas Theis, and Ferenc et al. Husza´r. Photo-realistic single image super-resolution using a generative adversarial network. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[56] Yulun Zhang, Kunpeng Li, Kai Li, Bineng Wang, Zhongyuan Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. *European Conference on Computer Vision (ECCV)*, 2018.

[57] Wenzhe Shi, Jose Caballero, Ferenc Husza´r, Johannes Totz, Andrew P. Aitken, and Robert Bishop. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[58] Jinsun Park, Dong-Won Kim, Jaehoon Lee, and Kwanghoon Sohn. Srfeat: Single image super-resolution with feature discrimination. *IEEE Access*, 2020.

[59] Kristina Simonyan and Stefan Fuertinger. Speech networks at rest and in action: interactions between functional brain networks controlling speech production. *Journal of Neurophysiology*, 113(7):2967–2978, 2015.

[60] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, and Andrew Y. Ng. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017.

[61] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2012.

[62] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[63] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[64] Muhammet Talo, Ugur B. Baloglu, Ozal Yildirim, and U. Rajendra Acharya. Automated classification of brain mri using convolutional neural networks.
*Pattern Recognition Letters*, 122:115–122, 2019.

[65] Chien-Yu Yeh et al. A hybrid deep learning model for brain tumor classification. *Frontiers in Neuroscience*, 14:924, 2020.

[66] Andre Esteva, Brett Kuprel, and Roberto A. et al. Novoa. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118, 2017.

[67] Geert Litjens, Thijs Kooi, Babak E. Bejnordi, Arnaud A. A. Setio, Francesco Ciompi, Mohsen Ghafoorian, and Clara I. Sa´nchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42:60–88, 2017.

[68] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *IEEE International Conference on Computer Vision (ICCV)*, 2017.

[69] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[70] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. In
*IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[71] Lutz Prechelt. Early stopping — but when? In Genevieve B. Orr and Klaus-Robert Mu¨ller, editors, *Neural Networks: Tricks of the Trade*, volume 1524 of *Lecture Notes in Computer Science*, pages 55–69, Berlin, Heidelberg, 1998. Springer.

[72] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision (ECCV)*, pages 694–711. Springer, 2016.

[73] P. Athisayamani et al. Brain tumor classification using resnet-152 with enhanced chimp optimization algorithm. *Computers in Biology and Medicine*, 152:106372, 2023.

[74] F. Zulfiqar and colleagues. Brain tumor classification using efficientnet and transfer learning. *Neural Computing and Applications*, 35:11789–11802, 2023.

[75] Jieneng Chen et al. Transunet: Transformers make strong

encoders for medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 485–495. Springer, 2021.

[76] Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R Roth, and Daguang Xu. Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. *Medical Image Analysis*, 75:102366, 2022.

[77] Asim Saeed, Muhammad Hussain, Imran Awan, and Sang Hyuk Lee. A gan-based hybrid deep learning model for improved brain tumor classification using mri images. *Biomedical Signal Processing and Control*, 73:103466, 2022.

[78] Umar Saeed, Abdul Rehman, Tanzila Saba, Amjad Rehman, Abdurrahman S. Bahaj, Muhammad Raza, and Maan Alqaysi. Enhanced deep convolutional neural network for brain tumor classification using mri images. *Computers in Biology and Medicine*, 144:105354, 2022.

[79]