# Multi-grade brain tumor classification using deep CNN with extensive data augmentation

Muhammad Sajjad [a,1], Salman Khan [b,1], Khan Muhammad [b], Wanqing Wu [c], Amin Ullah [b], Sung Wook Baik [b,*]

[a] Digital Image Processing Laboratory, Department of Computer Science, Islamia College, Peshawar, Pakistan
[b] Intellignet Media Laboratory, College of Electronics and Information Engineering, Sejong University, Seoul, Republic of Korea
[c] CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology (SIAT), Shenzhen, 518055, China

## ARTICLE INFO

## ABSTRACT

Numerous computer-aided diagnosis (CAD) systems have been recently presented in the history of medical imaging to assist radiologists about their patients. For full assistance of radiologists and better analysis of magnetic resonance imaging (MRI), multi-grade classification of brain tumor is an essential procedure. In this paper, we propose a novel convolutional neural network (CNN) based multi-grade brain tumor classification system. Firstly, tumor regions from an MR image are segmented using a deep learning technique. Secondly, extensive data augmentation is employed to effectively train the proposed system, avoiding the lack of data problem when dealing with MRI for multi-grade brain tumor classification. Finally, a pre-trained CNN model is fine-tuned using augmented data for brain tumor grade classification. The proposed system is experimentally evaluated on both augmented and original data and results show its convincing performance compared to existing methods.

© 2018 Published by Elsevier B.V.

## 1. Introduction

Brain tumor is one of the most dangerous and deadly cancer among adults and children. Early identification and classification of brain tumor into their specific grade is very important to treat the tumor effectively. According to the classification standards of World Health Organization (WHO), brain tumor or tumors of central nervous system is further divided into grades of malignancy grade I (benign) to grade IV (high malignancy) [1]. Glioblastomas (known as grade IV tumors by WHO) are considered the most lethal type of tumor carrying dismal prognosis even with many advancements in therapeutic management [2]. The primary method of differentiating grade IV tumors from other grades is through histopathology. Initial differentiation can be made on the basis of necrosis, microvascular proliferation, and vascular thrombosis features of grade IV tumors [3]. Conversely, these features are not generally transparent and might be hard to recognize, and pathologists are observed having assorted views about them [4].

Still there exist few drawbacks in the histopathological investigation, which makes the evaluation of gliomas unclear. To recognize types of glioma effectively without brain surgery which may threat human lives, MR imaging is introduced. MRI is used broadly since it has large variety of physiologically authoritative contrasts to recognize diverse tissues and enhances assessment of varied outlines of tissue configuration inside diffuse gliomas. In accumulation with typical sequences, several MRI techniques separate low-grade glioma from glioblastoma (grade IV) including diffusion-weighted imaging, MR spectroscopy, and perfusion-weighted imaging. A past review upheld that MRI scans are exceptionally particular for the treatment of brain stem gliomas and alternate to biopsies radiotherapy in many patients [5]. The role of MRI in diagnostic is very important that can lead to develop classification techniques for multi-grade brain tumor MRI data, resulting in fewer numbers of surgical operations.

CAD [6–8] or multi-grade brain tumor classification systems are established to support the radiologists in visualization [9] and defining tumor types [10–12]. With deep neural networks, high-level features extracted from the MRI scans are utilized to categorize different sorts of grades that might help radiologists in making decisions about early diagnosis and future treatment

---

approaches. This categorization presented by different researchers, provides the detail about malignancy of tumor by various percentages. The major problems with existing approaches for brain tumor classification are binary classification, lack of data samples for different grades, and low accuracy. The existing techniques classified brain tumor into only two classes (i.e., benign and malignant) which is very ambiguous for radiologist in further investigation or treatment. The next major challenge in brain tumor classification is deficiency of publicly available data. Moreover, the existing techniques till date have not shown any satisfying accuracy.

In this article, a novel deep learning-based framework is proposed to classify brain tumor grades, categorized by WHO standards. For the effective analysis of MRI images, first brain tumor is segmented from MRI. Next, the segmented tumor region within MRI is augmented using different augmentation techniques. Finally, the brain tumor is classified into four grades (i.e., grade I, grade II, grade III, and grade IV) by fine-tuning deep CNN model. The proposed system can be used as a second opinion for assistance of radiologist and can get recommendations of tumor grading after clinical examination of the patient. The key contributions of the paper can be summarized as follows:

1. To the best of our knowledge, this work is the first attempt toward the implementation of a CAD system for classification of brain tumor into four grades as per WHO standards for brain tumor classification.
2. The major problem in the field of MRI is lack of publicly available datasets. Although there are some datasets available on the Internet but the number of images specifically for our problem is very less. Thus, we use extensive data augmentation by augmenting the data with various parameters and different techniques to fill the gap of data and make the system transformational and noise invariant.
3. The CNN architecture employed in the proposed system has $3 \times 3$ kernels for all convolutional layers with 1 stride, unlike the other CNN architectures which have $11 \times 11$ or $7 \times 7$ kernels with 4–5 strides in the initial layers. Thus, we utilized small kernels CNN model, which is able to learn small texture patterns of multiple grades of tumor in MRI images.
4. Our system has achieved higher accuracy in brain MRI images compared to other state-of-the-art techniques despite the fact that some systems have achieved higher accuracy for processing pathological images.

The rest of this paper is organized as follows: In Section 2, background of multi-grade brain tumor classification is explained. Section 3 describes the proposed methodology with focus on segmentation, augmentation, and fine-tuning of the employed architecture. Section 4 presents the experimental tools and all the experiments before and after data augmentation. Finally, Section 5 concludes the paper with future work.

## 2. Background

In the field of medical image analysis a lot of work has been done and many researchers have contributed to various sub-fields of medical imaging [13,14]. This section explores past studies on brain tumor classification. In medical imaging most of the existing work refers to automatic segmentation of tumor region in MR images. Recently, many researchers have presented different techniques to detect and segment the tumor region in MR images [15–18]. After the tumor is segmented in MRI, it needs to be classified into different grades. In the early research studies [19–21] binary classifiers were used for various classes such as benign and malignant. For instance, Kharrat et al. [19] proposed a hybrid scheme for categoriz-

ing brain tumor into normal, benign, and malignant using genetic algorithm (GA) and support vector machine (SVM). Abdolmaleki et al. [20] developed a three-level backpropagation neural network to distinguish between malignant and benign tumors using thirteen different features. These features are selected on the basis of radiologists pictorial understanding. Their suggested method achieved classification accuracy of 91% and 94% for malignant and benign tumor, respectively, while performing experiments on the MRI data of 165 patients. Furthermore, Papageorgiou et al. [21] described brain tumors as low and high-grade gliomas based on fuzzy cognitive maps (FCM) in their work. FCM grading model comprises of 100 cases for experiments and attains an accuracy of 90.26% for low-grade and 93.22% for the high-grade brain tumors.

In addition to binary classification, researchers also have suggested multi-grade brain tumor classification [22–25]. For instance, Zacharaki et al. [22] presented a technique for classification of brain tumor into various types i.e., gliomas, primary gliomas, and metastases. This study comprises of numerous stages: region-of-interest (ROI) extraction followed by feature extraction, feature selection, and classification. Hsieh et al. [23] proposed a CAD system which extracts features from MRI data to measure the malignancy of gliomas. Performance of their method is evaluated on a database of 107 images, having 34 high-grade and 73 lower-grade gliomas images. The output accuracy of CAD system using global, local and fused feature are 76%, 83%, and 88%, respectively. Similarly, Sachdeva et al. [24] suggested a CAD system, in which they extracted texture and color features of segmented ROIs and selected optimal features through genetic algorithm (GA). The accuracy obtained from GA-SVM and GA-ANN is 91.7% and 94.9%, respectively. Furthermore, few researchers presented brain tumor classification methods, which focus on three types of brain tumors i.e., meningioma, glioma, and pituitary tumor. For instance, Cheng et al. [26] proposed a method with three different features extraction techniques, namely bag-of-words (BoW) model, intensity histogram, and gray level co-occurrence matrix (GLCM). Similarly, afshar et al. [27] presented a capsule networks (CapsNets) for brain tumor classification based on four objectives including adopt and incorporate CapsNets, investigation of over-fitting, capability of CapsNets, and development of visualization paradigm for output. Further, Zia et al. [28] suggested a generalized classification system for brain tumor based on rectangular window image cropping. Their system utilized discrete wavelet transform as a features extraction, PCA for dimensionality reduction, and SVM as a classifier.

Although there exist several techniques for brain tumor classification, but they have certain limitations which need to be considered while working with brain tumor segmentation and classification. The first limitation of these systems is their binary classification of tumor, which leaves many ambiguities for the radiologist. The reason is that classification into benign and malignant is not enough for radiologist to decide treatment and preventions for the patient. For clear and better understanding of radiologist, the classification needs to be multi-class, which classifies brain tumor into its relevant grades. Furthermore, lack of data is also a key challenge for researchers to achieve precise results. To address these key limitations, we proposed a deep learning-based framework with data augmentation technique to achieve effective results for multi-grade brain tumor classification.

## 3. Proposed method

Recently deep learning has been emerged due to its high rate of accuracy and its vast applications domain in many research areas including computer vision [29–31], image processing [32,33], authentication system [34], and speech recognition [35,36]. CNNs are feed-forward artificial neural networks (ANN) encouraged by
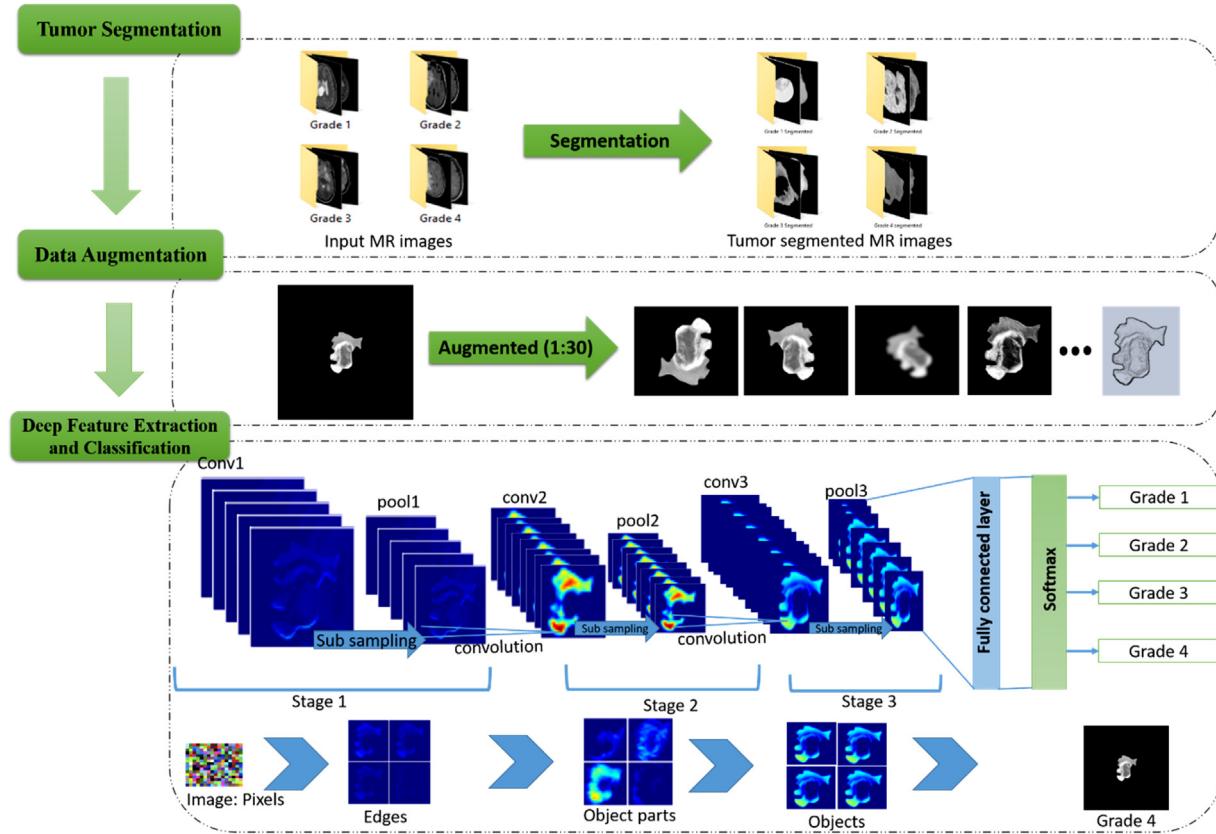
**Fig. 1.** CNN assisted multi-grade tumor segmentation and classification framework.

natural procedures intended to identify different patterns directly from image data. Motivated from the recent achievements of CNNs in various challenging tasks we used CNN to target the problem of multi-grade brain tumor classification. In this article, we present a novel deep learning framework to segment and classify brain tumor into four different grades using a fine-tuned CNN model. The proposed system consists of three main steps: 1) tumor segmentation, 2) data augmentation, and 3) deep features extraction and classification. In first step, tumor regions in both datasets are segmented. Tumor regions are segmented through a pre-trained CNN architecture whose layers explicitly support segmentation. Next step is data augmentation in which data is augmented through different transformational and noise invariance techniques by changing different parameters. In the third step, data is passed to a fine-tuned CNN architecture trained on brain MR images for the final prediction of tumor grades. The detailed proposed framework is shown in Fig. 1.

### 3.1. Tumor segmentation

In our proposed framework, the tumor is segmented using a fully automatic deep learning based technique called InputCascadeCNN [15]. A unique CNN architecture differs from other traditional CNNs due to its two-way processing of image. InputCascadeCNN uses a final layer that is a convolutional implementation of a fully connected layer which is 40-fold faster than the other state-of-the-art CNN models. The architecture of InputCascadeCNN consists of two streams: one with $7 \times 7$ accessible fields for extracting local features and another with $13 \times 13$ accessible fields for extracting global features. In the first step, by means of extracting the local features, the highest and lowest intensities are suppressed using a bias correction algorithm N4ITK [37]. In the next step, data of each input channel is normalized by subtracting mean channel and dividing it

by standard deviation of that channel to extract the global features. After applying these two steps, a post-processing step is applied to remove the noise from the segmented image. In post-processing, connected components labeling algorithm is applied to remove the flat blobs that may appear as a tumor region due to bright corners of the brains near the skull. As a result, the InputCascadeCNN model gives a segmented brain tumor region. The working procedure of InputCascadeCNN model for tumor segmentation is given in Fig. 2.

### 3.2. Data augmentation

In this section the process of data augmentation is discussed in detail. High-quality and abundant data is a key for effective deployment of various deep learning models [38]. The problem of classification addressed in this article lacks satisfactory amount of data to fed into deep learning architecture and acquire desired accuracy. Thus, to achieve the desired accuracy we extended the existing data by applying eight different augmentation techniques presented in Table 1. The possible employed augmentation techniques are rotation, flipping, skewness, and shears for geometric transformations invariance. The next four techniques such as Gaussian blur, sharpening, edge detection, and emboss are used for the noise invariance. In Table 1, second column refers to augmentation techniques along with their corresponding invariance parameters in third column. We have a total of 30 parameters and eight augmentation techniques, which extend each sample of the dataset into 30 samples.

### 3.3. Deep features extraction and classification

In the proposed system VGG-19 [39] CNN architecture is fine-tuned for brain tumor grades classification. VGG-19 architecture
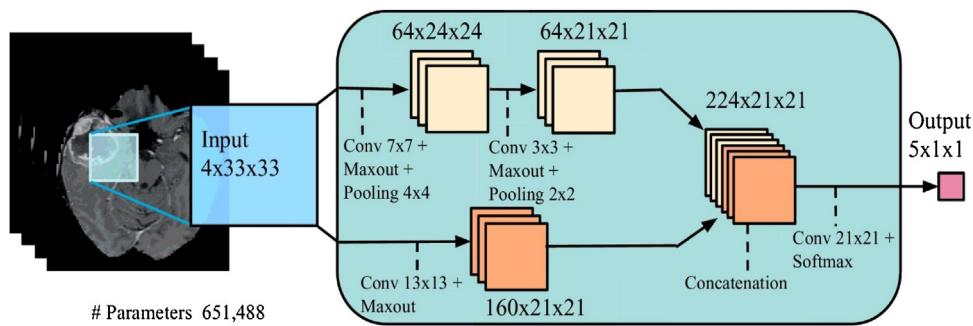
**Fig. 2.** InputCascadeCNN architecture for brain tumor segmentation through two way CNN.

**Table 1**
Different data augmentation techniques with their respective parameters.

| S.NO | Data augmentation technique | Parameters |
|------|------------------------------|------------|
| 1 | Rotation (Angle in degree) | −90° −45° 45° 90° |
| 2 | Flip | Top Bottom Right Left |
| 3 | Gaussian Blur (Sigma value) | 0.25 0.50 1.0 2.0 |
| 4 | Sharpen (lightness value) | 0.50 1.00 1.50 2.0 |
| 5 | Edges Detection (alpha value) | 0.25 0.50 0.75 1.00 |
| 6 | Emboss (strength value) | 0.50 1.00 1.50 2.0 |
| 7 | Skew (Tilt) | Left Right Forward Backward |
| 8 | Shear (axis & value) | X-axis 10 degrees Y-axis 10 degrees |

consists of 19 weighted layers, in which there are 16 convolutional and 3 fully connected (or dense) layers. In VGG-19 architecture, the first two convolutional layers are followed by max pooling, and the similar combination is repeated for subsequent two layers as presented in Fig. 3. The next eight layers are designed as a combination of four convolutional layers trailed by max pooling. Further, the last three layers are fully connected (FC), resulting in 4096, 4096, and 1000 features, respectively. The 1000-dimensional feature vector obtained from the last FC 8 layer is fed as an input to Softmax classifier, which makes the final prediction of brain tumor grade. The

motivation behind using VGG-19 is that there are $3 \times 3$ kernels for all the convolutional layers with 1 stride unlike other CNN models [40–42] which have $11 \times 11$ or $7 \times 7$ kernels with 4 to 5 strides in the initial layers. These types of large strides ignore the important patterns in the MRI and the higher size of kernels increases the number of parameters. We used the concept of transfer learning where the activation of pre-trained model are used to learn patterns in different data. Therefore, we have changed only the weights of fully connected layers for fine-tuning the same pre-trained model for brain tumor grades classification.

## 4. Experiments and results

In this section, we have discussed the experimental assessment of the proposed scheme for multi-grade brain tumor classification. We performed several experiments on two datasets including Radiopaedia [43] and brain tumor [26]. The datasets are divided into 50%, 25%, and 25% for training, cross validation, and testing sets, respectively. In the proposed system, we have increased the number of samples for each grade by utilizing different data augmentation techniques. Therefore, the evaluation of the proposed method with original and augmented data is discussed in separate sections. All experiments are evaluated using NVIDIA GTX-1070 with 8 GB onboard memory and deep learning framework Caffe [44] that are installed on Ubuntu 16.04.

### 4.1. Radiopaedia dataset

Radiopaedia [43] dataset consists of 121 MR images which are divided into four different grades as given in Table 2. The classification of the tumor is based on the ground truth labels given with each image according to the guidelines released in the yearly journal of WHO [1].

The number of images in each grade before data augmentation are very few, while deep learning-based technique requires a huge amount of data for effective patterns learning. Thus, by using data augmentation techniques, 30 new images are generated from each MR image. The number of newly created samples for each grade is given in Table 2.
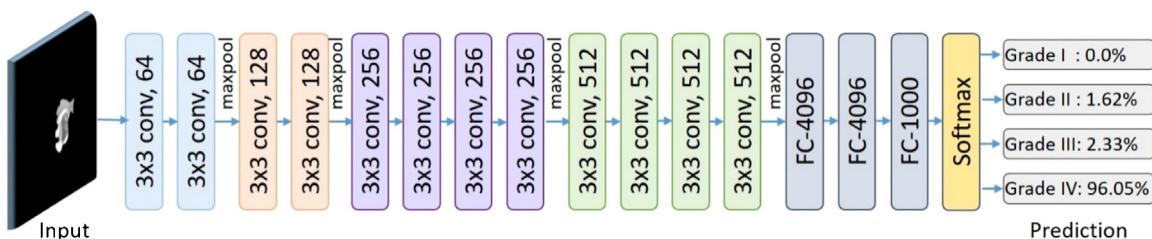


**Fig. 3.** Architecture of fine-tuned VGG-19 CNN model.

**Table 2**
Statistics of radiopaedia dataset before and after applying data augmentation.

| Tumor grades | Tumor type | Number of images | | Sub-types |
|---|---|---|---|---|
| | | Before augmentation | After augmentation | |
| Grade I | Menigiomas | 36 | 1080 | Craniopharnygioma, Pilocytic Astrocytoma, Pleomorphic xanthoastrocytoma |
| Grade II | Gliomas | 32 | 960 | Ependymoma, Low-grade Astrocytoma, and oligodendroglioma |
| Grade III | Gliomas | 25 | 750 | Anaplastic Astrocytoma and Anaplastic oligoastrocytoma |
| Grade IV | Glioblastmoas | 28 | 840 | Glioblastoma |

**Table 3**
Statistics of brain tumor dataset before and after applying data augmentation.

| Tumor types | Number of images | |
|---|---|---|
| | Before augmentation | After augmentation |
| Meningioma | 708 | 21240 |
| Glioma | 1426 | 42780 |
| Pituitary tumor | 930 | 27900 |

**Table 4**
Confusion matrix for test set of radiopaedia dataset before data augmentation.

| | Grade I (%) | Grade II (%) | Grade III (%) | Grade IV (%) |
|---|---|---|---|---|
| Grade I | **90.03** | 5.60 | 3.27 | 1.10 |
| Grade II | 4.56 | **89.91** | 5.53 | 0 |
| Grade III | 2.41 | 6.32 | **84.11** | 7.16 |
| Grade IV | 0 | 5.15 | 9.35 | **85.50** |
| Overall accuracy | **87.38%** | | | |

### 4.2. Brain tumor dataset

The brain tumor dataset [26] consists of 3064 T1-weighted contrast-enhanced MR images collected from 233 patients. Resolution of these images is $512 \times 512$ with pixel size of $0.49 \times 0.49$ mm$^2$. The gap between each slice is 1 mm and slice thickness is 6 mm. Statistics of this dataset are given in Table 3. The accuracy obtained on original dataset is improved after applying several data augmentation techniques as explained in Section 3.2. After data augmentation the number of samples in this dataset are increased from 3064 to 91,920 as given in Table 3.

### 4.3. Experimental results before data augmentation

The confusion matrices for radiopaedia and brain tumor datasets are given in Tables 4 and 5, respectively. It can be observed from Table 4 that the accuracy using radiopaedia dataset is 90.03%, 89.91%, 84.11%, and 85.50% for grade I, II, III, and IV, respectively. Although the results are not convincing enough, yet they can be trusted in real environment i.e., grade III is 7.16% confused with grade IV and grade IV is 9.35% confused with grade III. Further, the confusion matrix for brain tumor dataset using our proposed system before data augmentation along with sensitivity and specificity is given in Table 5. The overall sensitivity and specificity on this dataset is 84.51% and 93.34%, respectively.

### 4.4. Experimental results after data augmentation

In this section, the proposed system is evaluated after applying data augmentation over both datasets that made the results convincing enough to be considered in real scenarios. Confusion

**Table 6**
Confusion matrix of radiopaedia dataset after data augmentation.

| | Grade I (%) | Grade II (%) | Grade III (%) | Grade IV (%) |
|---|---|---|---|---|
| Grade I | **95.54** | 3.20 | 1.26 | 0 |
| Grade II | 2.80 | **92.66** | 4.54 | 0 |
| Grade III | 0.92 | 5.80 | **87.77** | 5.51 |
| Grade IV | 0 | 6.17 | 7.12 | **86.71** |
| Overall accuracy | **90.67%** | | | |

matrix of radiopaedia dataset with data augmentation is shown in Table 6. The accuracy using this dataset for grade I is increased from 90.03% to 95.5%. Similarly accuracies for grade II, III, and IV are increased from 89.91% to 92.66%, 84.11% to 87.77%, and 85.50% to 86.71%, respectively. The overall accuracy obtained is 90.67% which is much greater than accuracy achieved without data augmentation. On brain tumor dataset, data augmentation improved 4% sensitivity and 3% specificity, which increased the overall sensitivity to 88.41% and specificity to 96.12% as given in Table 7. Therefore, it is evident from the experiments that the data augmentation has a very positive impact on accuracy.

Training time validation accuracies for each epoch with and without data augmentation using radiopaedia dataset are visualized in Fig. 4. It is evident from Fig. 4 that abundant data improves the accuracy which proves that huge amount of data is essential for deep learning based techniques. For first epoch the accuracy before and after augmentation starts from 40% and 57%, respectively.

### 4.5. Comparison with state-of-the-art methods

A detailed comparison of the proposed method with state-of-the-art techniques is presented in Table 8. Papageorgiou et al. [21] classified brain tumor into low and high-grade on the basis of histopathological features and achieved an overall accuracy of 92%. Zacharaki et al. [22] achieved an accuracy of 80% by using only 102 images for classification of brain tumor into metastasis, grade II, III, and glioblastomas. Kharrat et al. [19] achieved an accuracy of 97% for binary classification of benign and malignant tumor with a total of 83 images in their dataset. An improvement in accuracy from 91.7 to 94% is achieved by Sachdeva et al. [24] due to the rich data of 428 images. Pinaya et al. [45] attained an accuracy of 73.6% using deep belief network and SVM. In another method, Barker et al. [46] classified pathological images into lower grade glioma and glioblastoma with an accuracy of 93.1% by utilizing spatially localized features with weighted voting for classification. A novel discriminative feature-oriented dictionary learning is used by Vu et al. [47] by classifying histopathological images with and without micro vascular proliferation and acquired an accuracy of 94.64%. Similarly, Hsieh et al. [23] utilized four global histogram along with

**Table 5**
Confusion matrix for test set of brain tumor dataset before data augmentation.

| | Meningioma(%) | Glioma(%) | Pituitary tumor(%) | Overall sensitivity(%) | Overall specificity (%) |
|---|---|---|---|---|---|
| Meningioma | 90.22 | 6.53 | 3.25 | | |
| Glioma | 4.19 | 93.12 | 2.69 | **84.51** | **93.34** |
| Pituitary tumor | 3.89 | 7.03 | 89.08 | | |

**Table 7**
Confusion matrix of brain tumor dataset after data augmentation.

|  | Meningioma(%) | Glioma(%) | Pituitary tumor(%) | Overall sensitivity(%) | Overall specificity (%) |
|---|---|---|---|---|---|
| Meningioma | 94.05 | 3.51 | 2.44 | | |
| Glioma | 2.83 | 96.14 | 10.3 | **88.41** | **96.12** |
| Pituitary tumor | 1.66 | 5.13 | 93.21 | | |

**Table 8**
Detailed comparison with state-of-the-art methods using radiopaedia dataset.

| S. No | Method / Year | Brain tumor classes and samples in each class / Total images | Features / Classifier | Class accuracy (%) | Overall accuracy (%) |
|---|---|---|---|---|---|
| 1 | Papageorgiou et al. [21] / 2008 | Low grade–41 High-grade–59 / 100 | Histopathological features, cellularity, mitoses, necrosis etc. / FCM | Low-grade–90.3 High-grade–93.2 | 92 |
| 2 | Zacharaki et al. [22] / 2009 | Metastasis–24 Meningiomas–4 Grade II–22 Grade III–18 Glioblastomas–34 / 102 | Tumor shape, Intensity, Rotational invariant texture features / SVM, ANN | Metastasis–91.7 Grade II–90.9 Grade III–33.4 Glioblastomas–41.2 | 80 |
| 3 | Kharrat et al. [19] / 2010 | Normal–29 Benign–32 Malignant–32 / 83 | Texture features using SGLDM / SVM | Normal–96 Benign–97 Malignant–97 | 94.44 |
| 4 | Sachdeva et al. [24] / 2016 | AS–118, GBM–59 MEN–97, MED–88, MET–66 / 428 | GLCM–16, LoG–4, DGTF–10, IBF–6, SBF–1, RILBPF–9 / SVM, ANN | AS–89.8 GBM–83.3 MED–95.6 MEN–91.8 MET–97.1 | 94 |
| 5 | Pinaya et al. [45] / 2016 | Healthy control–83, Schizophrenia–148 / 231 | DBN / SVM | Healthy control–76 Schizophrenia–70 | 73.6 |
| 6 | Barker et al. [46] / 2016 | Lower grade glioma–120, Glioblastoma multiforme–182 / 302 | Spatially localized features: Color, Shape, Textures features / weighted Voting. | Lower grade glioma–95 Glioblastoma multiforme–90 | 93.1 |
| 7 | Vu et al. [47] / 2016 | With MVP–57 Without MVP–133 / 190 | Discriminative Feature–oriented Dictionary Learning / (DFDL) | With MVP–94.59 Without MVP–94.69 | 94.64 |
| 8 | Hsieh et al. [23] / 2017 | Lower grade gliomas–73 Glioblastomas–34 / 107 | Four global histogram features, 14 local texture features / Logistic regression, ANN | Global features–76 Local features–83 | 88 |
| 9 | **Proposed method** | **Radiopaedia before augmentation** Grade I–36 Grade II–32 Grade III–25 Grade IV–28 / 121 **Radiopaedia after augmentation** Grade I–1080 Grade II–960 Grade III–750 Grade IV–840 / 3630 | Deep Convolution Neural Network Fine-tune VGG–19 / Softmax classifier | Grade I–90.03 Grade II–89.91 Grade III–84.11 Grade IV–85.50 Grade I–95.54 Grade II–92.66 Grade III–87.77 Grade IV–86.71 | **87.38** **90.67** |

14 local texture features and logistic regression for the classification of lower grade gliomas and glioblastomas.

The overall accuracy of their method is 76% and 83% for global and local features, respectively. The last row in Table 8 shows accuracy of the proposed system with and without data augmentation. The first experiment in last row indicates results before data augmentation where total number of images are 121 including 36, 32, 25, and 28 for Grade I, II, III, and IV, respectively. The proposed system achieved 90.03%, 89.91%, 84.11%, and 85.50% individual accuracy for grade I, II, III, and IV, respectively. The overall accuracy of our suggested technique is 87.38% for original data, which is increased to 90.67% after data augmentation.

Comparison using sensitivity, specificity, and accuracy with existing techniques using brain tumor dataset is presented in Table 9. Afshar et al. [27] evaluated their method on this dataset and achieved 86.56% accuracy. Zia et al. [27] proposed window based image cropping technique for brain tumor classification and achieved 86.26%, 90.90%, and 85.69% sensitivity, specificity, and accuracy, respectively. Similarly, Cheng et al. [26] obtained 81%, 92%, and 91.28% sensitivity, specificity, and accuracy, respectively. Our suggested approach overwhelmed the existing techniques and achieved the highest value of 88.41%, 96.12%, and 94.58% for sensitivity, specificity, and accuracy, respectively, which are the best results achieved so far over brain tumor dataset.
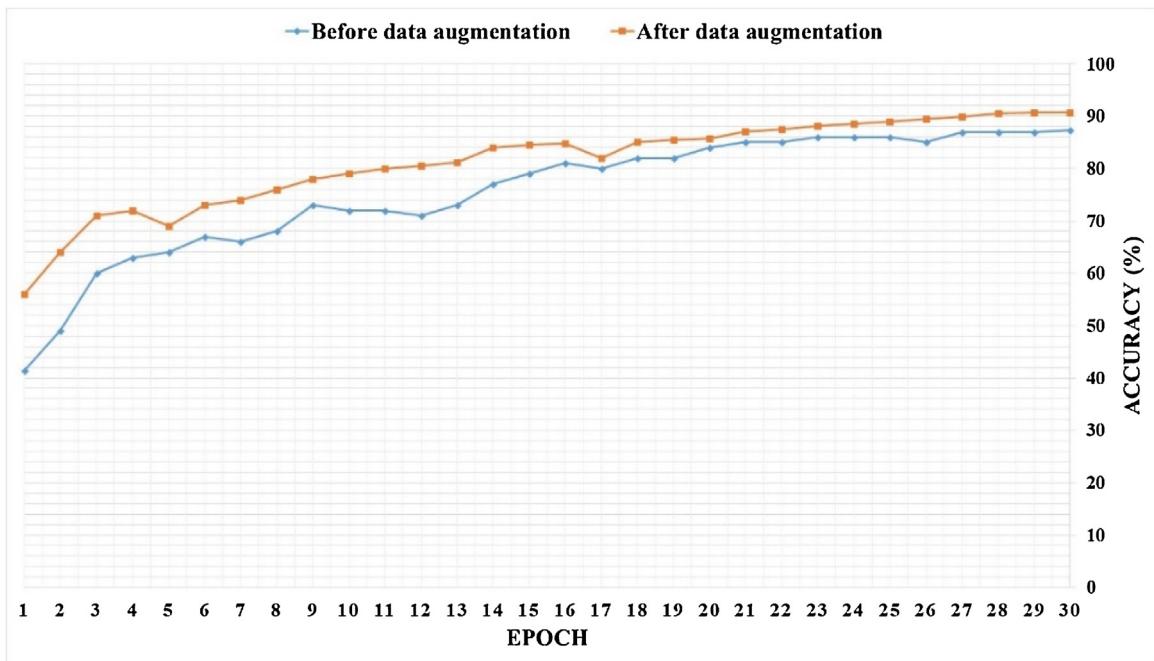
**Fig. 4.** Validation accuracy of the proposed fine-tuned CNN model before and after data augmentation for radiopaedia dataset.

**Table 9**
Comparison with state-of-the-art methods using brain tumor dataset.

| Method | Sensitivity(%) | Specificity(%) | Accuracy(%) |
|---|---|---|---|
| [27] | – | – | 86.56 |
| [28] | 86.26 | 90.90 | 85.69 |
| [26] | 81 | 92 | 91.28 |
| **Proposed** | **88.41** | **96.12** | **94.58** |

## 5. Conclusion and future work

In this paper, we presented a novel deep learning based multi-grade brain tumor classification system. Our system is threefold: 1) the tumor regions from the dataset are segmented through a CNN model, 2) the segmented data is further augmented using serval parameters to increase the number of data samples, and 3) a pre-trained VGG-19 CNN model is fine-tuned for multi-grade brain tumor classification. In the proposed system we improved the accuracy by utilizing data augmentation and deep learning. The experimental results show the effectiveness of the proposed CNN based CAD system, for helping the radiologist in taking a precise decision for classifying multi-grade brain tumors into four grades. In the future, we aim to extend our current work for fine-grained classification of each grade with investigation of light-weight CNN architectures to balance the efficiency and accuracy.

## Acknowledgment

## References

[1] D.N. Louis, A. Perry, G. Reifenberger, A. Von Deimling, D. Figarella-Branger, W.K. Cavenee, et al., The 2016 World Health Organization classification of tumors of the central nervous system: a summary, Acta Neuropathol. 131 (2016) 803–820.

[2] E.G. Van Meir, C.G. Hadjipanayis, A.D. Norden, H.K. Shu, P.Y. Wen, J.J. Olson, Exciting new advances in neuro-oncology: the avenue to a cure for malignant glioma, CA. Cancer J. Clin. 60 (2010) 166–193.

[3] D.N. Louis, H. Ohgaki, O.D. Wiestler, W.K. Cavenee, P.C. Burger, A. Jouvet, et al., The 2007 WHO classification of tumours of the central nervous system, Acta Neuropathol. 114 (2007) 97–109.

[4] C.L. Nutt, D. Mani, R.A. Betensky, P. Tamayo, J.G. Cairncross, C. Ladd, et al., Gene expression-based classification of malignant gliomas correlates better with survival than histological classification, Cancer Res. 63 (2003) 1602–1607.

[5] A.L. Albright, R.J. Packer, R. Zimmerman, L.B. Rorke, J. Boyett, G.D. Hammond, Magnetic resonance scans should replace biopsies for the diagnosis of diffuse brain stem gliomas: a report from the Children's Cancer Group, Neurosurgery 33 (1993) 1026–1030.

[6] M. Sajjad, S. Khan, Z. Jan, K. Muhammad, H. Moon, J.T. Kwak, et al., Leukocytes classification and segmentation in microscopic blood smear: a resource-aware healthcare service in smart cities, IEEE Access 5 (2017) 3475–3489.

[7] Z. Jan, A. Khan, M. Sajjad, K. Muhammad, S. Rho, I. Mehmood, A review on automated diagnosis of malaria parasite in microscopic blood smears images, Multimedia Tools Appl. 77 (2018) 9801–9826.

[8] M. Elhoseny, A. Abdelaziz, A.S. Salama, A.M. Riad, K. Muhammad, A.K. Sangaiah, A hybrid model of internet of things and cloud computing to manage big data in health services applications, Future Gener. Comput. Syst. 86 (2018) 1383–1394.

[9] I. Mehmood, M. Sajjad, K. Muhammad, S.I.A. Shah, A.K. Sangaiah, M. Shoaib, et al., An efficient computerized decision support system for the analysis and 3D visualization of brain tumor, Multimedia Tools Appl. (2018) 1–26.

[10] K.L.-C. Hsieh, R.-J. Tsai, Y.-C. Teng, C.-M. Lo, Effect of a computer-aided diagnosis system on radiologists' performance in grading gliomas with MRI, PLoS One 12 (2017) e0171342.

[11] G.J. Litjens, J.O. Barentsz, N. Karssemeijer, H.J. Huisman, Clinical evaluation of a computer-aided diagnosis system for determining cancer aggressiveness in prostate MRI, Eur. Radiol. 25 (2015) 3187–3199.

[12] S.-H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, H. Cheng, Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling, J. Med. Syst. 42 (2018) 85.

[13] S.-H. Wang, K. Muhammad, P. Phillips, Z. Dong, Y.-D. Zhang, Ductal carcinoma in situ detection in breast thermography by extreme learning machine and combination of statistical measure and fractal dimension, J. Ambient Intell. Hum. Comput. (2017) 1–11.

[14] N.C. Codella, D. Gutman, M.E. Celebi, B. Helba, M.A. Marchetti, S.W. Dusza, et al., Skin lesion analysis toward melanoma detection, in: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted by the International Skin Imaging Collaboration (ISIC),Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, 2018, pp. 168–172.

[15] M. Havaei, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, et al., Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.

[16] B.H. Menze, A. Jakab, S. Bauer, J. Kalpathy-Cramer, K. Farahani, J. Kirby, et al., The multimodal brain tumor image segmentation benchmark (BRATS), IEEE Trans. Med. Imaging 34 (1993) 2015.

[17] M. Prastawa, E. Bullitt, N. Moon, K. Van Leemput, G. Gerig, Automatic brain tumor segmentation by subject specific modification of atlas priors 1, Acad. Radiol. 10 (2003) 1341–1348.

[18] T. Ateeq, M.N. Majeed, S.M. Anwar, M. Maqsood, Z.-u. Rehman, J.W. Lee, et al., Ensemble-classifiers-assisted detection of cerebral microbleeds in brain MRI, Comput. Electric. Eng. (2018).

[19] A. Kharrat, K. Gasmi, M.B. Messaoud, N. Benamrane, M. Abid, A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine, Leonardo J. Sci. 17 (2010) 71–82.

[20] P. Abdolmaleki, F. Mihara, K. Masuda, L.D. Buadu, Neural networks analysis of astrocytic gliomas from MRI appearances, Cancer Lett. 118 (1997) 69–78.

[21] E. Papageorgiou, P. Spyridonos, D.T. Glotsos, C.D. Stylios, P. Ravazoula, G. Nikiforidis, et al., Brain tumor characterization using the soft computing technique of fuzzy cognitive maps, Appl. Soft Comput. 8 (2008) 820–828.

[22] E.I. Zacharaki, S. Wang, S. Chawla, D. Soo Yoo, R. Wolf, E.R. Melhem, et al., Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme, Magn. Reson. Med. 62 (2009) 1609–1618.

[23] K.L.-C. Hsieh, C.-M. Lo, C.-J. Hsiao, Computer-aided grading of gliomas based on local and global MRI features, Comput. Methods Programs Biomed. 139 (2017) 31–38.

[24] J. Sachdeva, V. Kumar, I. Gupta, N. Khandelwal, C.K. Ahuja, A package-SFERCB-"Segmentation, feature extraction, reduction and classification analysis by both SVM and ANN for brain tumors", Appl. Soft Comput. 47 (2016) 151–167.

[25] W. Jia, K. Muhammad, S.-H. Wang, Y.-D. Zhang, Five-category classification of pathological brain images based on deep stacked sparse autoencoder, Multimedia Tools Appl. (2017) 1–20.

[26] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, et al., Enhanced performance of brain tumor classification via tumor region augmentation and partition, PloS one 10 (2015) e0140381.

[27] P. Afshar, A. Mohammadi, and K. N. Plataniotis, Brain tumor type classification via capsule networks, ärXiv preprint arXiv:1802.10200, 2018.

[28] R. Zia, P. Akhtar, A. Aziz, A new rectangular window based image cropping method for generalization of brain neoplasm classification systems, Int. J. Imaging Syst. Technol. 28 (2018) 153–162.

[29] K. Muhammad, J. Ahmad, Z. Lv, P. Bellavista, P. Yang, S.W. Baik, Efficient deep CNN-based fire detection and localization in video surveillance applications, in: IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, pp. 1–16.

[30] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, S.W. Baik, Action recognition in video sequences using deep Bi-directional LSTM with CNN features, IEEE Access 6 (2018) 1155–1166.

[31] K. Muhammad, J. Ahmad, S.W. Baik, Early fire detection using convolutional neural networks during surveillance for effective disaster management, Neurocomputing 288 (2018) 30–42.

[32] J. Ahmad, K. Muhammad, J. Lloret, S.W. Baik, Efficient conversion of deep features to compact binary codes using fourier decomposition for multimedia Big data, IEEE Trans. Ind. Inf. 14 (2018) 3205–3215.

[33] J. Ahmad, K. Muhammad, S.W. Baik, Medical image retrieval with compact binary codes generated in frequency domain using highly reactive convolutional features, J. Med. Syst. 42 (2018) 24.

[34] M. Sajjad, S. Khan, T. Hussain, K. Muhammad, A.K. Sangaiah, A. Castiglione, et al., CNN-based anti-spoofing two-tier multi-factor authentication system, Pattern Recognit. Lett. (2018).

[35] J. Ahmad, K. Muhammad, S.-i. Kwon, S.W. Baik, S. Rho, Ďempster-shafer fusion based gender recognition for speech analysis applications, in: Platform Technology and Service (PlatCon), 2016 International Conference on, 2016, pp. 1–4.

[36] A.M. Badshah, N. Rahim, N. Ullah, J. Ahmad, K. Muhammad, M.Y. Lee, et al., Deep features-based speech emotion recognition for smart affective services, Multimedia Tools Appl. (2017) 1–19.

[37] B.B. Avants, N. Tustison, G. Song, Advanced normalization tools (ANTS), Insight J. 2 (2009) 1–35.

[38] Y.-D. Zhang, Z. Dong, X. Chen, W. Jia, S. Du, K. Muhammad, et al., Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation, Multimedia Tools Appl. (2017) 1–20.

[39] K. Simonyan and A. Zisserman, V̈ery deep convolutional networks for large-scale image recognition,ärXiv preprint arXiv: 1409.1556, 2014.

[40] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.

[41] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, Överfeat: Integrated recognition, localization and detection using convolutional networks,ärXiv preprint arXiv:1312.6229, 2013.

[42] M.D. Zeiler, R. Fergus, Visualizing and understanding convolutional networks, in: European Conference on Computer Vision, 2014, pp. 818–833.

[43] Radiopaedia. https://radiopaedia.org/.

[44] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., Caffe: convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 675–678.

[45] W.H. Pinaya, A. Gadelha, O.M. Doyle, C. Noto, A. Zugman, Q. Cordeiro, et al., Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia, Sci. Rep. 6 (2016) 38897.

[46] J. Barker, A. Hoogi, A. Depeursinge, D.L. Rubin, Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles, Med. Image Anal. 30 (2016) 60–71.

[47] T.H. Vu, H.S. Mousavi, V. Monga, G. Rao, U.A. Rao, Histopathological image classification using discriminative feature-oriented dictionary learning, IEEE Trans. Med. Imaging 35 (2016) 738–751.

**Muhammad Sajjad** received his master's degree from Department of Computer Science, College of Signals, National University of Sciences and Technology, Rawalpindi, Pakistan. He received his PhD degree in Digital Contents from Sejong University, Seoul, Republic of Korea. He is now working as an assistant professor at Department of Computer Science, Islamia College Peshawar, Pakistan. He is also head of "Digital Image Processing Laboratory (DIP Lab)" at Islamia College Peshawar, Pakistan., where students are working on research projects such social data analysis, medical image analysis, multi-modal data mining and summarization, image/video prioritization and ranking, Fog computing, Internet of Things, virtual reality, and image/video retrieval under his supervision. His primary research interests include computer vision, image understanding, pattern recognition, and robot vision and multimedia applications, with current emphasis on raspberry-pi and deep learning-based bioinformatics, video scene understanding, activity analysis, Fog computing, Internet of Things, and real-time tracking.

**Salman Khan** received his bachelor's degree in computer science from Islamia College Peshawar, Pakistan in 2017 with research in medical image analysis. Currently, he is pursuing MS degree in software convergence from Sejong University, Seoul, Republic of Korea. He is working as a research associate at Intelligent Media Laboratory (IM Lab). His research interests include feature extraction, medical image analysis, content-based image retrieval, computer vision and deep learning.

**Khan Muhammad** (S'16-M'18) received the Ph.D degree in Digital Contents from Sejong University, South Korea. He is currently working as a researcher at Intelligent Media Laboratory (IM Lab), Sejong University, Seoul. His research interests include medical image analysis (brain MRI, diagnostic hysteroscopy and wireless capsule endoscopy), information security (steganography, encryption, watermarking and image hashing), video summarization, computer vision, fire/smoke scene analysis, and video surveillance. He has published over 50 papers in peer-reviewed international journals and conferences in these research areas with target venues as IEEE COMMAG, TII, TIE, TSMC-Systems, Access, TSC, Elsevier INS, Neurocomputing, PRL, FGCS, COMCOM, COMIND, JPDC, PMC, BSPC, CAEE, Springer NCAA, MTAP, JOMS, and RTIP, etc. He is also serving as a professional reviewer for over 40 well-reputed journals and conferences.

**Dr. Wanqing Wu** received his B.S. degree in computer science and technology from Hunan Normal University, Hunan, China, in 2004, an ME in computer science from Chongqing University, Chongqing, China, in 2007, and a Ph.D. from the Pusan National University of Computer Engineering, Korea, in 2013. He is currently an Associate Professor at the Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China. He also works within the Key Laboratory for Health Informatics of the Chinese Academy of Sciences, China. His research interests include wearable sensing technology, biomedical signal sensing and processing, body sensor networks, biofeedback, biometric security, internet of things and multimodal medical data fusion. Dr. Wu has published over 40 papers in renowned journals and conferences, mostly in the area of wearable sensor technology, wireless body sensor network and Internet of Medical Things, and has also contributed to the peer review of 10 journals. He has secured over 15 major competitive research grants.

**Amin Ullah** received the BS degree in computer science from the Islamia College Peshawar, Peshawar, Pakistan. He is currently pursuing the M.S. leading to the Ph.D. degree with the Intelligent Media Laboratory, Sejong University, South Korea. His research interests include human actions and activity recognition, sequence learning, image and video analysis, and deep learning for multimedia understanding.

**Sung Wook Baik** received the B.S degree in computer science from Seoul National University, Seoul, Korea, in 1987, the M.S. degree in computer science from Northern Illinois University, Dekalb, in 1992, and the Ph.D. degree in information technology engineering from George Mason University, Fairfax, VA, in 1999. He worked at Datamat Systems Research Inc. as a senior scientist of the Intelligent Systems Group from 1997 to 2002. In 2002, he joined the faculty of the College of Electronics and Information Engineering, Sejong University, Seoul, Korea, where he is currently a Full Professor and Dean of Digital Contents. He is also the head of Intelligent Media Laboratory (IM Lab) at Sejong University. He served as professional reviewer for several well-reputed journals such as IEEE Communication Magazine, Sensors, Information Fusion, Information Sciences, IEEE TIP, MBEC, MTAP, SIVP and JVCI. His research interests include computer vision, multimedia, pattern recognition, machine learning, data mining, virtual reality, and computer games. He is a professional member of the IEEE.