PENet: Prior evidence deep neural network for bladder cancer staging

Xiaoqian Zhou, Xiaodong Yue, Zhikang Xu, Thierry Denoeux, Yufei Chen

PENet: Prior Evidence Deep Neural Network for Bladder Cancer Staging

Xiaoqian Zhou[a], Xiaodong Yue[a,b,*], Zhikang Xu[a], Thierry Denoeux[c,d], Yufei Chen[e]

[a]*School of Computer Engineering and Science, Shanghai University, Shanghai, China*
[b]*Artificial Intelligence Institute of Shanghai University, Shanghai, China*
[c]*Sino-European School of Technology, Shanghai University, Shanghai, China*
[d]*Université de technologie de Compiégne, Compiégne, France*
[e]*College of Electronics and Information Engineering, Tongji University, Shanghai, China*

Abstract

Bladder cancer is a heterogeneous, complicated, and widespread illness with high rates of morbidity, death, and expense if not treated adequately. The accurate and exact stage of bladder cancer is fundamental for treatment choices and prognostic forecasts, as indicated by convincing evidence from randomized trials. The extraordinary capability of Deep Convolutional Neural Networks (DCNNs) to extract features is one of the primary advantages offered by these types of networks. DCNNs work well in numerous real clinical medical applications as it demands costly large-scale data annotation. However, a lack of background information hinders its effectiveness and interpretability. Clinicians identify the stage of a tumor by evaluating whether the tumor is muscle-invasive, as shown in images by the tumor's infiltration of the bladder wall. Incorporating this clinical knowledge in DCNN has the ability to enhance the performance of bladder cancer staging and bring the prediction into accordance with medical principles. Therefore, we introduce PENet, innovative prior evidence deep neural network, for classifying MR images of bladder cancer staging in line with clinical knowledge. To do this, first, the degree to which the tumor has penetrated into the bladder wall is measured to get prior distribution parameters of class probability called

*Corresponding author
Email addresses: zhouxiaoqian@shu.edu.cn (Xiaoqian Zhou), yswantfly@shu.edu.cn (Xiaodong Yue), xuzhikangnba@shu.edu.cn (Zhikang Xu), thierry.denoeux@utc.fr (Thierry Denoeux), yufeichen@tongji.edu.cn (Yufei Chen)

prior evidence. Second, we formulate the posterior distribution of class probability according to Bayesian Theorem. Last, we modify the loss function based on posterior distribution of class probability which parameters include both prior evidence and prediction evidence in the learning procedure. Our investigation reveals that the prediction error and the variance of PENet may be reduced by giving the network prior evidence that is consistent with the ground truth. Using MR image datasets, experiments show that PENet performs better than image-based DCNN algorithms for bladder cancer staging.

Keywords: Bladder cancer staging, deep neural network, prior evidence fusion

1. Introduction

Bladder cancer is a varied, complicated, and widespread illness with a changeable natural history that, if not treated adequately, is linked with high morbidity, fatality rates, and expense[1, 2]. Bladder cancer continues to evolve into two distinct subtypes, producing non-muscle-invasive tumors and non-papillary muscle-invasive tumors[3]. Complete excision of the tumor is the most frequent treatment for non-muscle-invasive bladder cancer (NMIBC), while radical cystectomy and neoadjuvant chemotherapy are the most common curative procedures for muscle-invasive bladder cancer (MIBC)[1]. Currently, the TNM staging system is the most prevalent tumor staging system and the standard approach for doctors to stage malignant tumors. NMIBC and MIBC correlate to T1 or less and T2 or more for bladder cancer T staging. Therefore, accurate and exact bladder cancer staging is extremely important, since it influences the treatment strategy and prognosis. Radiographs, such as those produced by magnetic resonance imaging (MRI) and computed tomography (CT), are the primary diagnostic tools and treatment modalities used in clinical settings to diagnose and treat of bladder cancer. As is well-known, DCNN has strong feature extraction capabilities and a promising future in clinical medical imaging, such as computer-aided diagnosis (CAD) of bladder cancer based on images[4], which covers cancer staging[5, 6], tumor segmentation[7, 8, 9], cancer treatments[10, 11], among others. Most known image-based approaches for staging bladder cancer rely on direct DCNNs to assess the stage of tumors based on images of patients' bladders. However, pure image-based classification methods necessitate costly large-scale data annotation as a foundation and disregard clinical

experiences and priors, which may lead to inaccurate tumor stage predictions inconsistent with medical knowledge, thereby limiting performance and interpretability. In contrast, clinical physicians determine the stage of a bladder tumor by detecting whether or not the tumor has spread into the surrounding muscle, which is also referred to as tumor infiltration into the bladder wall. Particularly effective for enhancing bladder cancer staging predictions and bringing them into compliance with medical law is incorporating clinical priors of tumor infiltration into DCNN. In order to achieve the stage prediction of bladder cancer in line with clinical knowledge, we introduce a novel prior evidence deep neural network called PENet.

First, by measuring how much a bladder tumor has infiltrated the bladder wall, we may use evidence theory[12, 13] to estimate how much evidence there is to support our hypothesis. So the prior distribution parameters of class probability are obtained called prior evidence. Second, Bayesian Theorem are used to directly formulate posterior distribution of class probability by likelihood distribution from observation and prior distribution of class probability. Last, we reconstructed the objective function of the evidential deep neural network based on posterior distribution of class probability to improve the performance of PENet by integrating prior evidence guidance. The contributions of this study are outlined in the following paragraph.

- *Propose a straightforward strategy for quantifying clinical features of tumor penetration into the bladder wall to obtain prior distribution parameter.* It is possible to determine how much overlap exists between the tumor and the bladder wall by determining how many matrices of the tumor and the bladder wall are in the inner product of each other to generate prior evidence as prior distribution parameter.

- *Construct posterior distribution of class probability.* We formulate posterior distribution based on Bayesian Theorem according to likelihood distribution from observation and prior distribution where distribution parameters are extracted evidences from images and prior respectively.

- *Propose a PENet for classifying bladder cancer stage images.* We reformulate the objective function of evidential deep neural networks for PENet based on posterior distribution of class probability to optimize PENet's weight parameters. As we proved, PENet's prediction errors may be reduced when fused prior evidence compatible with ground truth.

The remaining sections are grouped as follows. In Section 2, we discuss the relevant work. Section 3 discusses our proposed DCNN-based bladder cancer staging technique PENet, including the representation of clinical prior evidence of bladder tumor infiltration and the approach for fusing the prior evidence into DCNN for bladder cancer staging. Section 4 contains the experimental data that support the efficiency of PENet in assessing the bladder cancer stage. Conclusion is presented in Section 5.

2. Related Work

2.1. Computer-aided bladder cancer diagnosis

Bladder cancer, one of the most frequent malignant tumors of the urinary system[14, 15], with significant rates of morbidity, mortality, and cost. Depending on the proper stage of bladder cancer, treatment options and expected prognosis will change[16]. The TNM staging system is currently the most common tumor staging system and is the standard method for clinicians to classify malignant tumors. A partial or total cystectomy is often used to treat tumors of T2 or higher (MIBC), which may be diagnosed by MR imaging and staged from T0 to T4 depending on whether the tumor is muscle-invasive or the degree of infiltration. DCNNs have seen widespread application as a method of computer-aided detection for bladder cancer due to their capacity to automatically extract hierarchical characteristics from images at varying degrees of image abstraction, which allows DCNNs to analyze images at varying levels of detail [17, 18]. Applications encompass bladder segmentation[19, 20],tumor detection[21], cancer therapies, and bladder cancer staging[5].

For bladder segmentation, Ma et al. suggested an automatic bladder segmentation approach based on U-Net for CT urography, where the bladder boundaries do not need user input for the bounding box and are estimated by U-Net[7]. Shkolyar et al. suggested CystoNet with DCNN to improve the performance of bladder tumor localization, surgical resection, and intraoperative navigation for tumor detection[22]. For cancer treatments, Rundo et al. developed a non-invasive prediction system comprised of a CT scan image pipeline and a radionics pipeline to characterize the expected response to immunotherapy, therefore informing physicians of treatment alternatives[23].

Characteristics were automatically retrieved from medical pictures to stage the bladder tumors using DCNNs for bladder cancer staging. [6] improved cancer staging prediction by combining morphological and textural

features specific to bladder staging with many classifiers, including support vector machines, neural networks, and random forests. Functional characteristics representing percentiles of the cumulative distribution function (CDF), morphological features representing radionics texture features, and morphological features defining tumor shape were retrieved from T2W-MRI, and DW-MRI as input to neural networks for bladder cancer staging [24]. Zhang et al. learned infiltration criteria from MR images that are advantageous for tumor staging based on clinical experiences and used the rules into DCNN to increase performance[25]. Using ResNet structure, non-local attention, and image super-resolution processing, [26] developed a model with high performance for CT imaging-based bladder cancer staging.

2.2. Evidence-theory-driven machine learning

Evidence theory (Dempster-Shafer evidence theory) is viewed as a type of generalized probability that applies Dempster's rule to reasoning while using the mass function to evaluate decision-making uncertainty[12, 13]. Through quantifying views and their uncertainty, evidence theory has been extensively employed in the areas of information fusion and reasoning with uncertainty. By merging evidence theory with machine learning, a number of supervised and unsupervised learning methods for uncertain data analysis have been developed, including evidence K-nearest neighbor[27], evidence linear discrimination analysis[28], and neural network with evidence[29].

A combination of evidence theory with medical image analysis, such as medical picture segmentation, has been made. By employing Dempster's rule to combine data from each voxel's vicinity in MR images, Capelle et al. suggested a region-based segmentation approach for brain tumor segmentation[30]. According to Lian et al., the problem of unclear and imprecise segmentation in each modality may be addressed by applying Dempster's rule to fuse the results of distinct modality segmentation[31]. Belief function and Dempster's rule are used by Huang et al. to measure the uncertainty of segmentation in the border area and increase performance[32]. For the diagnosis of pneumonia from chest X-ray pictures, [33] applies the Dempster-Shafer theory to the fusion of five pre-trained convolutional neural networks, including VGG16, Xception, InceptionV3, ResNet50, and DenseNet201, and offers good detection performance. Evidence theory is used for a variety of additional activities in the medical image analysis process in addition to being used for the segmentation of medical images.

In order to increase the effectiveness of evidence theory in processing complex data, several researchers have recently merged it with deep learning techniques to create evidential deep neural networks[34] that construct classification uncertainty[35]. Unlike conventional deep neural networks, evidential deep neural networks treat the activation values of the output layer as evidences retrieved from the data for prediction, and the prediction of the network is therefore extended to a probability distribution with evidence parameters. As a result, evidential deep neural networks provide a method for calculating the degree of uncertainty in deep neural network predictions and correcting the inaccurate ones [36, 37].

3. Method

3.1. Workflow

Two major components of the proposed PENet are the measurement of prior clinical experiences for bladder cancer staging and formulation of posterior probability distribution of class probability based on Bayesian Theory for bladder cancer staging. Three elements of PENet's process are shown in Figure 1. The first highlighted module with a green dashed line depicts the process of creating tumor stage evidences (When we want to express two types of evidence for high or low cancer stage, or two sources of evidence for data and clinical experiences, we use the word "evidences".) from labeled MR images using a deep convolutional neural network. The second module, shown by an orange dashed line, depicts the procedure for retrieving prior evidence of tumor staging from segmentation masks. The third module, shown by the purple dotted line, we can formulate likelihood distribution from observation using extracted evidence from images(viewed as random variable) as parameter; Similarly, based on extracted prior evidence, we formulate prior distribution of class probability. Through Bayesian Theorem, the posterior distribution of class probability can be derived with two sources evidence parameters which is the basis of loss function for PENet to improve performance.

3.2. Evaluation of tumor infiltration's clinical experiences

The stages of bladder cancer may be categorized into five phases, from T0 to T4, as mentioned in Section 1. Cystectomy is necessary if the stage is higher than T2, which is considered a high stage[3]. The degree of tumor infiltration into the bladder wall is a common way for human physicians to

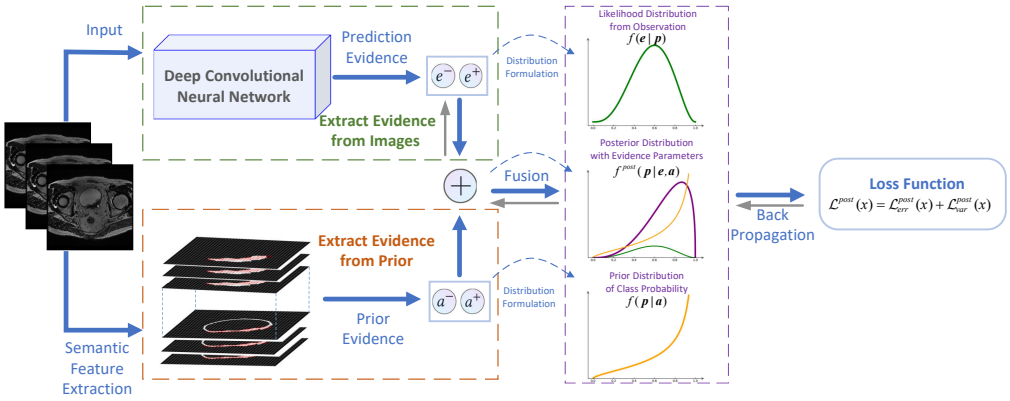

Figure 1: Workflow of PENet.

determine the stage of bladder tumors, based on clinical experiences. Therefore, high tumor-bladder wall overlap suggests high cancer stage ($\geq$ T2), whereas little overlap indicates low stage ($<$ T2). Since tumor and bladder wall segmentation masks are used in clinical previous staging of bladder cancer, we determined the degree of overlap

$$\varrho = \frac{\langle M_{tumor}, M_{wall} \rangle}{\|M_{wall}\|_2^2}, \tag{1}$$

assuming that M_{tumor} and M_{wall} refer to the MR image-derived tumor and bladder wall mask matrices, respectively, $\langle \cdot, \cdot \rangle$ represents the Frobenius inner product of the two. Obtain ϱ_i by further normalizing the overlap degree ρ_i for each picture i, such as $\rho_i = \frac{\varrho_i - \varrho^{min}}{\varrho^{max} - \varrho^{min}}$.

Using two instances, shown in Figure 2, we can better understand how much overlap there is between tumors and the bladder wall and how advanced the disease is. When ρ rises, so does the stage of bladder cancer, which may range anywhere from $<$ T2 to $\geq$ T2. In the following part, we used ρ to create prior evidence to fuse into PENet for bladder cancer stage.

3.3. PENet to classify bladder cancer stage

The deep neural network can be understood as stacking several non-linear function with a softmax operator on top to discriminate the training data, which is parameter regression framework of Multinomial distribution. Currently, we are all aware that the traditional deep neural network is overconfident since the denominator of softmax has squished the output probabilities,

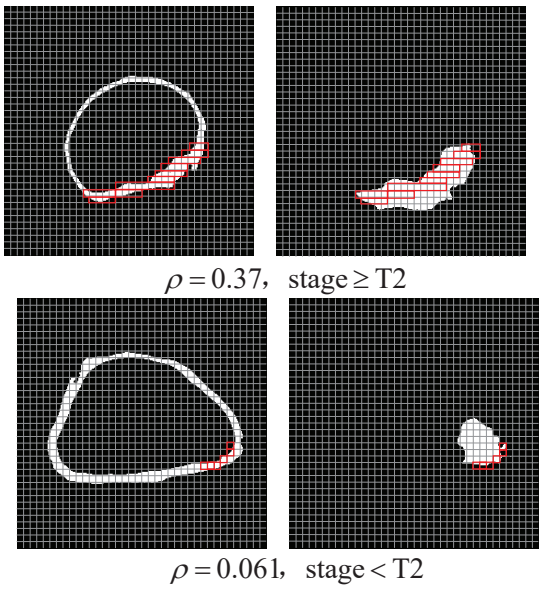

Figure 2: Calculating prior evidences from segmentation of bladder tumor and wall.

and the point estimation of class probability which is first-order uncertainty can be proceed by the cross entropy loss function but cannot express the variance of prediction probability such as second-order uncertainty. As mentioned in Section 2.2, evidential deep neural network (EvidentialNet)[34] proposes a principled way to formulate the prediction probability distribution with evidence parameter which can directly express the variance of prediction and accomplish reliable classification. In contrast to EvidentialNet, we consider the fact that the class probability's prior distribution should not be disregarded when class probability is seen as a random variable. Moreover, we can directly infer the expression of the posterior distribution of the class probability through the distribution assumption and Bayesian Theorem. Through the loss function reconstructed by the posterior distribution, PENet can provide more stable and accurate prediction to improve performance.

In this paragraph, we will carefully derive the expression for the posterior distribution of class probability. First, according to the observations of the DCNN, we assume that the likelihood function is a *Binomial* distribution $f(\boldsymbol{e} \mid \boldsymbol{p}) = C^{e^+}_{e^+ + e^-} p_0^{e^-} p_1^{e^+}$, where distribution parameter $\boldsymbol{e} = (e^-, e^+)$ is extracted evidence from DCNN to support stage $<$ T2 and stage $\geq$ T2 respectively, and $\boldsymbol{p} = (p_0, p_1)$ is class probability. Second, on the basis of the

conjugate prior distributions, we assume that the prior distribution about class probability is a $Beta$ distribution $f(\boldsymbol{p} \mid \boldsymbol{a}) = Beta(\boldsymbol{p} \mid 2a^-, 2a^+) = \frac{1}{B(2a^-,2a^+)} p_0^{2a^--1} p_1^{2a^+-1}$, where $\boldsymbol{a} = (a^-, a^+)$ is the prior evidence of tumor infiltration that support stage $<$ T2 and stage $\geq$ T2 respectively. This prior distribution function satisfies that when there is no prior knowledge for classification $a^- = a^+ = 0.5$, we have $Beta(\boldsymbol{p} \mid 2a^-, 2a^+) = 1$, which is a uniform distribution. So we can directly conclude that the posterior distribution $f^{post}(\boldsymbol{p} \mid \boldsymbol{e}, \boldsymbol{a})$ is also a $Beta$ distribution. According to Bayesian theorem, we can directly calculate the specific expression of the posterior probability distribution as

$$
\begin{aligned}
f^{post}(\boldsymbol{p} \mid \boldsymbol{e}, \boldsymbol{a}) &= \frac{f(\boldsymbol{p} \mid \boldsymbol{a}) * f(\boldsymbol{e} \mid \boldsymbol{p})}{\int_0^1 f(\boldsymbol{p} \mid \boldsymbol{a}) * f(\boldsymbol{e} \mid \boldsymbol{p}) d\boldsymbol{p}}, \\
&= \frac{C_{e^+ + e^-}^{e^+} p_0^{e^-} p_1^{e^+} * \frac{1}{B(2a^-,2a^+)} p_0^{2a^--1} p_1^{2a^+-1}}{\int_0^1 C_{e^+ + e^-}^{e^+} p_0^{e^-} p_1^{e^+} * \frac{1}{B(2a^-,2a^+)} p_0^{2a^--1} p_1^{2a^+-1} d\boldsymbol{p}}, \\
&= \frac{p_0^{e^-+2a^--1} p_1^{e^++2a^+-1}}{\int_0^1 p_0^{e^-+2a^--1} p_1^{e^+2a^+-1} d\boldsymbol{p}}, \\
&= \frac{p_0^{e^-+2a^--1} p_1^{e^++2a^+-1}}{\mathcal{B}(e^- + 2a^-, e^+ + 2a^+)}, \\
&= Beta(e^- + 2a^-, e^+ + 2a^+), \qquad (2)
\end{aligned}
$$

where $\mathcal{B}(\cdot)$ is $Beta$ function. We are able to construct the prior evidences a_i^-, a_i^+ for each samples i as the parameters of f^{post} by

$$
a_i^- = 1 - \rho_i, \ a_i^+ = \rho_i. \qquad (3)
$$

which is based on the value of ρ_i that was acquired by quantifying the clinical prior in Section 3.2 for each sample.

Given $\mathcal{D}=\{x_i, \boldsymbol{y}_i\}_{i=1}^N$ of N labeled MR images, is subject to the following declarations: one-hot label vector x_i, $\boldsymbol{y}_i = (0, 1)$ when stage $\geq$ T2, and $\boldsymbol{y}_i = (1, 0)$ when stage $<$ T2, are represented by $\boldsymbol{y}_i = (y_{i0}, y_{i1})$. We are able to define the loss function as the expectation of mean squares error $\|\boldsymbol{y}_i - \boldsymbol{p}_i\|_2^2$ on the basis of $f^{post}(\boldsymbol{p} \mid \boldsymbol{e}, \boldsymbol{a})$. The loss function of PENet is

denoted as $\sum_{i=1}^{N} \mathcal{L}^{post}(x_i)$, for each sample x_i the loss is

$$\begin{aligned}
\mathcal{L}^{post}(x_i) &= \int \|\boldsymbol{y}_i - \boldsymbol{p}_i\|_2^2 f^{post}(\boldsymbol{p}_i \mid \boldsymbol{e}_i, \boldsymbol{a}_i) d\boldsymbol{p}_i \\
&= E_{f^{post}}\left[\|\boldsymbol{y}_i - \boldsymbol{p}_i\|_2^2\right] \\
&= \sum_{j=0}^{1} E_{f^{post}}[y_{ij}^2 - 2y_{ij}p_{ij} + p_{ij}^2] \\
&= \sum_{j=0}^{1} (y_{ij}^2 - 2y_{ij}^2 E_{f^{post}}(p_{ij}) + E_{f^{post}}(p_{ij}^2)).
\end{aligned} \tag{4}$$

Due to $E_{f^{post}}(p_{ij}^2) = E_{f^{post}}(p_{ij})^2 + Var_{f^{post}}(p_{ij})$, we are able to deduce the equation and arrive at

$$\begin{aligned}
\mathcal{L}^{post}(x_i) &= \sum_{j=0}^{1} \left(y_{ij} - E_{f^{post}}(p_{ij})\right)^2 + Var_{f^{post}}(p_{ij}) \\
&= \underbrace{\left(y_{i0} - \frac{e_i^- + 2a_i^-}{\mathcal{S}_i}\right)^2 + \left(y_{i1} - \frac{e_i^+ + 2a_i^+}{\mathcal{S}_i}\right)^2}_{\mathcal{L}_{err}^{post}} \\
&\quad + \underbrace{\frac{2(e_i^- + 2a_i^-)(e_i^+ + 2a_i^+)}{\mathcal{S}_i^2(\mathcal{S}_i + 1)}}_{\mathcal{L}_{var}^{post}},
\end{aligned} \tag{5}$$

where $\mathcal{S}_i = e_i^- + e_i^+ + 2$. Because of this novel loss function, we are able to evaluate whether or not we can improve the performance for bladder cancer stage. We can deduce from Equation (5) that the loss function of PENet is made up of two components: the prediction error term $\mathcal{L}_{err}^{post}$ and the prediction variance term $\mathcal{L}_{var}^{post}$ respectively. This indicates that we may reduce the prediction error as well as the variance in the data at the same time by reducing the loss function concurrently. In addition, we make the assumption that the prediction error and variance are denoted by the symbols $\mathcal{L}_{err}$ and $\mathcal{L}_{var}$, respectively, when class probability prior distribution is a uniform distribution. Through the application of the following two theorems, we conduct an investigation into how the accuracy of the prediction may be improved by posterior distribution of class probability with evidence parameters.

Theorem 1. *When $a^- < a^+$ in a negative case or $a^+ > a^-$ in a positive case, we have $\mathcal{L}_{err} > \mathcal{L}_{err}^{post}$.*

proof: The prediction error term $\mathcal{L}_{err}$ of EvidentialNet may be obtained as follows with uniform prior distribution

$$\mathcal{L}_{err} = \left(y_0 - \frac{e^- + 1}{\mathcal{S}}\right)^2 + \left(y_1 - \frac{e^+ + 1}{\mathcal{S}}\right)^2. \tag{6}$$

On the other hand, predictive error terms given by posterior distribution is

$$\mathcal{L}_{err}^{post} = \left(y_0 - \frac{e^- + 2a^-}{\mathcal{S}}\right)^2 + \left(y_1 - \frac{e^+ + 2a^+}{\mathcal{S}}\right)^2. \tag{7}$$

Due to the fact that $a^+ > a^-, a^+ + a^- = 1$, we may deduce that $e^- + 2a^- < e^- + 1$ and $e^+ + 2a^+ > e^+ + 1$ respectively. Further inference may be drawn from the fact that (6) and (7) both indicate

$$\begin{aligned} \left(y_0 - \frac{e^- + 1}{\mathcal{S}}\right)^2 &> \left(y_0 - \frac{e^- + 2a^-}{\mathcal{S}}\right)^2, \\ \left(y_1 - \frac{e^+ + 1}{\mathcal{S}}\right)^2 &> \left(y_1 - \frac{e^+ + 2a^+}{\mathcal{S}}\right)^2, \\ \mathcal{L}_{err} &> \mathcal{L}_{err}^{post}. \end{aligned} \tag{8}$$

The procedure for providing evidence for negative cases is the same as described previously.

Theorem 2. *We have $\mathcal{L}_{var} > \mathcal{L}_{var}^{post}$ if $(e^+ - e^-)(a^+ - a^-) > 0$ in each and every case.*

proof: The prediction variance term $\mathcal{L}_{var}$ of EvidentialNet may be obtained as follows with uniform prior distribution

$$\mathcal{L}_{var} = \frac{2(e^- + 1)(e^+ + 1)}{\mathcal{S}^2(\mathcal{S} + 1)}. \tag{9}$$

We have shown that the prediction variance term fused prior distribution $\mathcal{L}_{var}^{post}$ in (5) gets the following results:

$$\begin{aligned} \mathcal{L}_{var} - \mathcal{L}_{var}^{post} &= \frac{2(e^- + 1)(e^+ + 1)}{\mathcal{S}^2(\mathcal{S} + 1)} - \frac{2(e^- + 2a^-)(e^+ + 2a^+)}{\mathcal{S}^2(\mathcal{S} + 1)} \\ &= \frac{e^+(1 - 2a^-) + e^-(1 - 2a^+) + (1 - 4a^+a^-)}{\mathcal{S}^2(\mathcal{S} + 1)}. \end{aligned}$$

As a result that $(1 - 4a^+a^-) \geq 0$ and $a^+ + a^- = 1$, we obtain

$$\mathcal{L}_{var} - \mathcal{L}_{var}^{post} = \frac{(e^+ - e^-)(a^+ - a^-) + (1 - 4a^+a^-)}{\mathcal{S}^2(\mathcal{S}+1)} > 0. \tag{10}$$

With the help of Theorem 1's derivation proof, we can create an approach for figuring out whether or not prior evidence may be incorporated into PENet and so selecting out evidence that could increase prediction error. When the prior evidence is consistent with the ground truth, the posterior distribution calculated by the integrated prior distribution can reduce the prediction error of PENet. Regarding Theorem 2, we are aware that if the evidences of the classification and the prior evidences are consistent, PENet can provide lower prediction variance by a sharp posterior distribution of class probability. It is our goal to use the Bayesian theory to create the *Beta* posterior distribution of class probability with two sources evidence parameters and reformulate the objective function as the expectation of prediction error, which may be reduced by the derivation proofs of the prediction error and variance. Algorithm 1 provides a summary of the process of model training.

***Algorithm* 1** Workflow of model training

1. Calculate prior evidence $\boldsymbol{a}$ from segmentation of bladder tumor and wall according to (1) and (4);
2. Generate prediction evidence $\boldsymbol{e}$ from images through deep neural network;
3. According to (3), compute posterior distribution of class probability and obtain $\mathcal{L}^{post}(x_i)$ for each images;
4. Optimize the PENet by decreasing $\mathcal{L}_{err}^{post}$ and $\mathcal{L}_{var}^{post}$ until convergence.

4. Experiment Results

During the course of the experiments, we gather the MR images of patients with bladder cancer for stage prediction from two distinct parts: our cooperation hospital and the Chinese University Computer Design Competition. The dimension of each MR image is 512 pixels on 512 pixels, and the image data collection comprises 344 T2-weighted MR images of 38 individuals. High stage to low stage ratio is 1.26 to 1, and all MR images have

been identified as MIBC (stage $\geq$ T2) or NMIBC by human clinicians (stage <T2).

In order to apply classification algorithms to MR images and accurately forecast the cancer stage (high or low), we divided the data set by patients and carried out five-fold cross validation in order to put our experiments into action. Metrics such as *accuracy*, *precision*, *recall rate* and $F1$-*score* are also used to assess the effectiveness of image classification techniques. In all, there are two parts to the experiments. The first part of the analysis will determine whether the planned PENet is successful in predicting the stage of bladder cancer. PENet's superiority over other representative picture classification algorithms will be tested in the second part.

4.1. Increasing efficiency of predictions by prior integration

Ablation experiments are used in this part to verify the prediction improvement obtained by incorporating clinical prior evidence into DCNN. In order to stage bladder cancer, we make use of three different deep convolutional neural networks, namely ResNet[38], EvidentialNet[34], and PENet. ResNet18's backbone network model [38] is used in all of the DCNNs mentioned. Table 1 provides an overview of the classification performance. When we compare ResNet with EvidentialNet, we discover that evidence theory's capacity to turn prediction probabilities into distributions of prediction probabilities leads to performance improvements. All of the assessment metrics of accuracy, precision, recall rate, and F1−score are further raised by +8.91%, +11.44%, +4.96%, and +8.35% respectively when the stable posterior distribution is obtained by PENet by superimposing prior evidences to EvidentialNet. PENet's increased capacity to stage bladder cancer was shown by this experiment, which indicated that prior evidence integration improved the prediction of the disease.

Table 1: Ablation studies of integrating prior evidences

	ResNet	EvidentialNet	PENet	Improvement
Accuracy	72.14	83.86	**92.77**	+8.91%
Precision	70.43	80.40	**91.84**	+11.44%
Recall	84.19	90.05	**95.01**	+4.96%
F1-score	76.06	84.84	**93.19**	+8.35%

Next, some hard-to-classified cases which are shown in Figure 3 are provided to verify the performance improvement of PENet over EvidentialNet due to integrating prior distribution with prior evidence parameter and we will analyze it from three aspects.

The first aspect is that PENet can rectify samples that are misclassified by EvidentialNet, as shown in the first four cases in Figure 3. For low stage,The shape and intensity of the shaded region within the red circle correspond to bladder tumors which is why EvidentialNet produced inaccurate stage $\geq$T2 forecasts. The correct classification was accomplished in Figure 3(a) by integrating the evidence of non-tumor-infiltration into PENet to enhance the prediction probability to 0.86. In a similar way, PENet was able to enhance the low-stage predictions given by EvidentialNet at $p_0 = 0.24$ to $p_0 = 0.77$ for proper classification in Figure 3(c).This is the result of making the class probability posterior distribution closer to the true distribution by learning the prior evidence from the segmentation mask.

For high stage, EvidentialNet may come to the wrong conclusion about the stage of the cancer if the intensity of the tumor is comparable to the bladder wall. In Figure 3(b) and (d), the quantified previous evidences of tumor-infiltration made it possible for PENet to enhance the predictions given by EvidentialNet from $p_1 = 0.35$ and $p_1 = 0.50$ to $p_1 = 0.93$ and $p_1 = 0.76$, respectively.

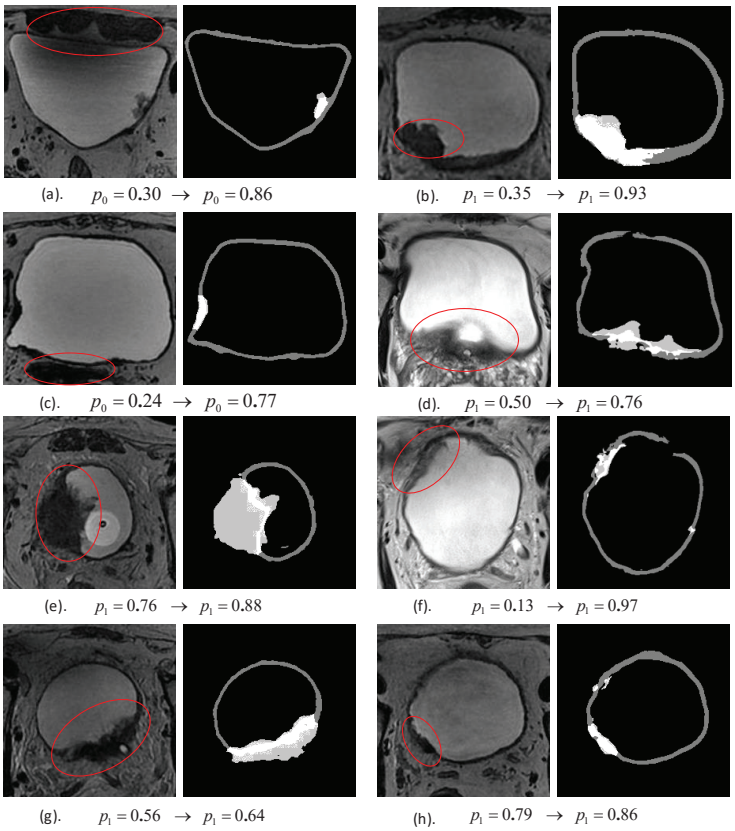

Figure 3: Cases of cancer stage prediction improvements brought by prior evidences.

The second aspect is that fusing prior evidences into PENet can enhance the confidence of correct prediction produced by EvidentialNet. In Figure 3(e), PENet improves the prediction of EvidentialNet from $p_1 = 0.76$ to $p_1 = 0.88$, which improves the confidence of the prediction due to prior evidences. Figure 3(g) provides a case of a low confidence prediction for EvidentialNet. PENet improves confidence by boosting prediction from $p_1 = 0.56$ to $p_1 = 0.64$.

The third aspect is that the prior evidences can help PENet to deal with the confusing cases. Figure 3(f) and (h) present two cases that are very challenging for EvidentialNet to correctly identify the tumor stage. EvidentialNet fails to recognize the existence of the tumor in Figure 3(f) because the tumor is so intimately linked to the wall of the bladder. As a result, the EvidentialNet generates an incorrect prediction of $p_1 = 0.13$. On the other

hand, in our approach, precise prior evidence parameter directs PENet to generate the correct prediction of high cancer stage $p_1 = 0.97$. The tumor's attachment to the bladder wall is quite thin in Figure 3(h). EvidentialNet, while its proper classification, may not concentrate on the lesion's precise location. According to the PENet's improved prediction of $p_1 = 0.86$, a more advanced stage of cancer was confirmed.

Extracting clinical prior evidence to get reliable posterior distribution of class probability for PENet provides more effective correction of misclassifications, increased confidence in prediction, and improved capacity to handle the hard instances. Because of this, PENet is a promising tool for integrating clinical evidence to improve performance and interpretation.

4.2. Compared to different bladder cancer staging methods

We tested PENet's performance against that of four other DCNN-based image classification techniques such as ResNet18[38], DenseNet[39], EvientialNet[34], and a rule-integrated approach such as RuleNet[25]. We discovered that PENet's performance was superior to each of these other methods. In Figure 4 and Table 2, the evaluation metrics generated by the five different approaches are shown, while Table 3 displays the more refined findings of the five-fold cross-validation.

Table 2: Comparison of different cancer stage classification methods.

Methods	Accuracy	Precision	Recall	F1 score
ResNet	72.14	70.43	84.19	76.06
DenseNet	78.80	75.43	87.19	80.87
EvidentialNet	83.86	80.84	90.05	84.84
RuleNet	85.24	83.68	91.08	86.80
PENet	**92.77**	**91.84**	**95.01**	**93.19**

Because integrating prior evidence parameter makes PENet's prediction of the cancer stage more accurate and trustworthy, we may find that PENet improves all assessment metrics in comparison to other methods. Because integrating prior evidences helps PENet better predict the cancer stage with reliable posterior distribution. RuleNet and PENet, both of which are directed by clinical rules and prior evidences, outperform solely data-driven DCNN models such as ResNet, DenseNet, and EvidentialNet, as can be shown in the Figure 4. On the other hand, RuleNet takes a lot of data for

Table 3: Cross validation results of cancer stage classification methods.

Cross Validation 1	Accuracy	Precision	Recall	F1 score
ResNet	85.06	82.19	**100.0**	90.23
DenseNet	85.06	87.30	91.67	89.43
EvidentialNet	89.66	86.96	**100.0**	93.02
RuleNet	82.76	80	**100.0**	88.88
Our Method	**93.10**	**93.55**	96.67	**95.08**
Cross Validation 2	Accuracy	Precision	Recall	F1 score
ResNet	84.13	85.19	79.31	82.14
DenseNet	77.78	72.73	**82.76**	77.42
EvidentialNet	84.13	82.76	**82.76**	82.76
RuleNet	88.88	90.24	85.86	88.00
Our Method	**90.48**	**96.00**	**82.76**	**88.89**
Cross Validation 3	Accuracy	Precision	Recall	F1 score
ResNet	69.12	69.44	71.43	70.42
DenseNet	80.88	72.92	**100.0**	84.34
EvidentialNet	83.82	81.58	88.57	84.93
RuleNet	83.82	77.27	97.14	86.07
PENet	**92.65**	**87.50**	**100.0**	**93.33**
Cross Validation 4	Accuracy	Precision	Recall	F1 score
ResNet	61.36	53.57	78.95	63.83
DenseNet	75.00	68.18	78.95	73.17
EvidentialNet	77.27	71.43	78.95	75.00
RuleNet	86.36	88.23	78.94	83.33
PENet	**95.45**	**90.48**	**100.0**	**95.00**
Cross Validation 5	Accuracy	Precision	Recall	F1 score
ResNet	61.04	61.76	91.3	73.68
DenseNet	75.32	77.55	82.61	80.00
EvidentialNet	84.42	79.43	**100.0**	88.46
RuleNet	84.41	82.69	93.47	87.75
PENet	**92.21**	**91.67**	95.65	**93.62**

an additional network to acquire clinical decision rules, whereas PENet performs better on a less amount of training data. In addition, prior evidences are chosen in accordance with Theorems 1 and 2 to guarantee that the integrated prior evidences will minimize prediction error and variance, which ultimately leads to more accurate and stable predictions.

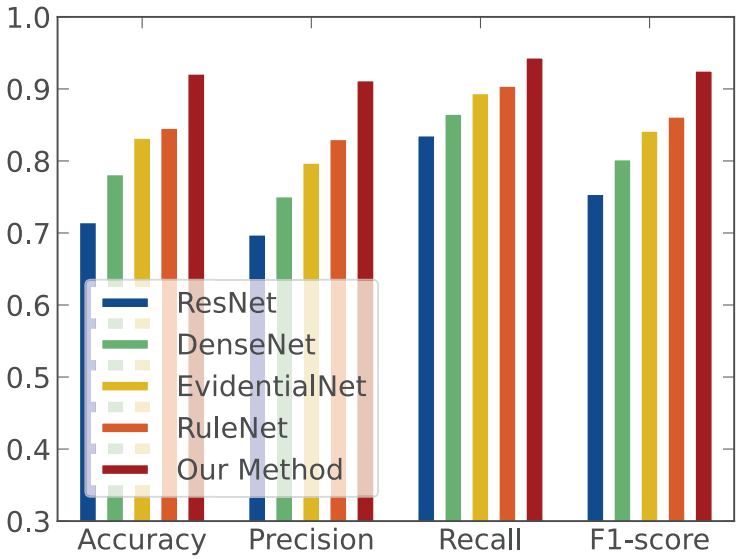

Figure 4: Comparison of different cancer stage classification methods.

Last, the regions of interest (ROI) in MR images that were created by five different classification methods were shown in an intelligible way by Grad-CAM[40]. Each of the ROI visualizations that have been developed is shown in the Figure 5.

It has been observed that the distribution of ROIs produced by ResNet and DenseNet in MR images is unstable and that it comprises numerous areas that are unrelated to bladder cancers. Possible causes include the fact that the limited labeled MR images do not guarantee that ResNet and DenseNet concentrate on the lesion region where the tumor is situated and that they might be impacted by areas that are not associated. In contrast to these two methods, EvidentialNet may partly relieve the data shortages by utilizing evidence to formalize the prediction distribution in order to accomplish classification. According to EvidentialNet's prediction performance, it is not steady even though the shapes and intensity of the areas are similar to bladder tumor. Clinical experience rules included into RuleNet provide more

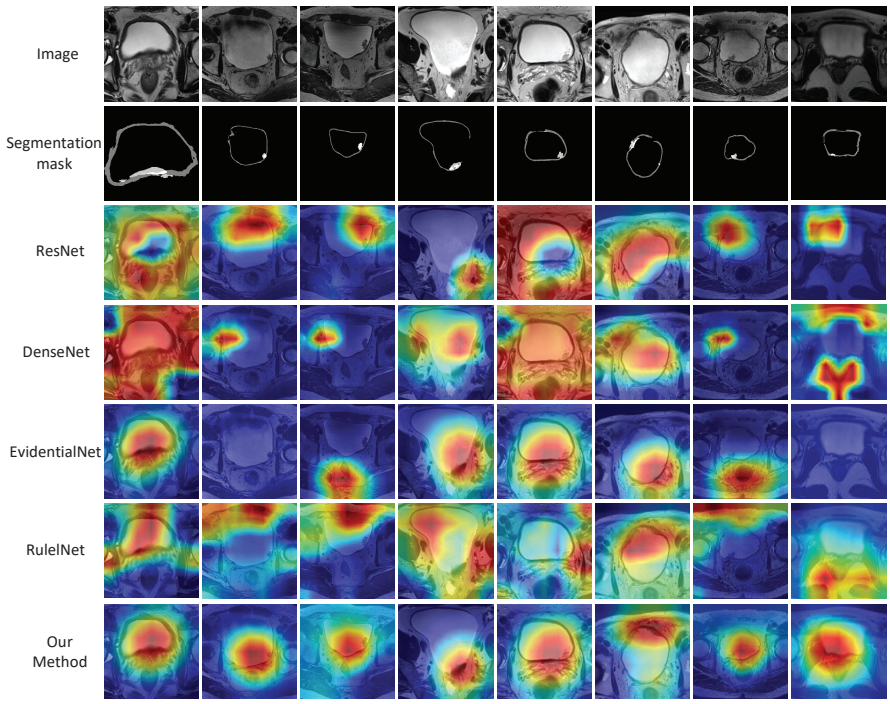

Figure 5: Visualization of ROIs in MR images produced by comparative classification methods.

accurate and stable ROIs than data-driven DCNNs but they are sensitive in the background noise area. With the help of incorporating prior evidences from clinical experiences that are extracted from the tumor-wall segmentation masks to get reliable posterior distribution, PENet's ROIs focus on the overlap regions between the bladder tumor and the wall stably and robustly. In conclusion, as shown by the outcomes of ROI visualization, the capacity of PENet to make precise predictions about the stage of cancer is beneficial from its capacity to integrate knowledge gained from clinical experiences.

5. Conclusion

It is possible that deep neural networks cannot accurately anticipate the bladder cancer stage at which cancer will be found due to a lack of relevant clinical expertise. We offer a strategy that is effective and for generating prior evidences of tumor-infiltration in order to quantify clinical experiences. Following that, we fused prior evidences as posterior distribution parameter

of class probability into PENet in order to increase the accuracy of cancer stage prediction. Experiments have shown that PENet is beneficial for identifying the stage of bladder cancer. In future research, a lot of focus will be taken at the gap that may be seen between the predictions offered by neural networks and the clinical knowledge that has already been gathered.

Acknowledgement

This work was supported by National Natural Science Foundation of China (Serial Nos. 62173252, 61976134), Natural Science Foundation of Shanghai (NO. 21ZR1423900) and Open Project Foundation of Intelligent Information Processing Key Laboratory of Shanxi Province, China (No. CICIP2021001).

References

[1] A. M. Kamat, N. M. Hahn, J. A. Efstathiou, S. P. Lerner, P.-U. Malmström, W. Choi, C. C. Guo, Y. Lotan, W. Kassouf, Bladder cancer, The Lancet 355 388 (10061) (2016) 2796–2810.

[2] Z. Kirkali, T. Chan, M. Manoharan, F. Algaba, C. Busch, L. Cheng, L. Kiemeney, M. Kriegmair, R. Montironi, W. M. Murphy, et al., Bladder cancer: epidemiology, staging and grading, and diagnosis, Urology 66 (6) (2005) 4–34.

[3] O. Sanli, J. Dobruch, M. A. Knowles, M. Burger, M. Alemozaffar, M. E. Nielsen, Y. Lotan, Bladder cancer, Nature reviews Disease primers 3 (1) (2017) 1–19.

[4] M. E. Gosnell, D. M. Polikarpov, E. M. Goldys, A. V. Zvyagin, D. A. Gillatt, Computer-assisted cystoscopy diagnosis of bladder cancer, in: Urologic Oncology: Seminars and Original Investigations 36 (2018) 8-e9.

[5] S. S. Garapati, L. M. Hadjiiski, K. H. Cha, H.-P. Chan, E. M. Caoili, R. H. Cohan, A. Weizer, A. Alva, C. Paramagul, J. Wei, et al., Automatic staging of bladder cancer on ct urography, in: Medical Imaging 2016: Computer- Aided Diagnosis 9785 (2016) 367–372.

[6] S. S. Garapati, L. Hadjiiski, K. H. Cha, H.-P. Chan, E. M. Caoili, R. H. Cohan, A. Weizer, A. Alva, C. Paramagul, J. Wei, et al., Urinary bladder cancer staging in ct urography using machine learning, Medical physics 44 (11) (2017) 5814–5823.

[7] X. Ma, L. M. Hadjiiski, J. Wei, H.-P. Chan, K. H. Cha, R. H. Cohan, E. M. Caoili, R. Samala, C. Zhou, Y. Lu, U-net based deep learning bladder segmentation in ct urography, Medical physics 46 (4) (2019) 1752–1765.

[8] K. H. Cha, L. Hadjiiski, R. K. Samala, H.-P. Chan, E. M. Caoili, R. H. Cohan, Urinary bladder segmentation in ct urography using deep-learning convolutional neural network and level sets, Medical physics 43 (4) (2016) 1882–1896.

[9] X. Xu, F. Zhou, B. Liu, Automatic bladder segmentation from ct images using deep cnn and 3d fully connected crf-rnn, International journal of computer assisted radiology and surgery 13 (7) (2018) 967–975.

[10] K. H. Cha, L. Hadjiiski, H.-P. Chan, A. Z. Weizer, A. Alva, R. H. Cohan, E. M. Caoili, C. Paramagul, R. K. Samala, Bladder cancer treatment response assessment in ct using radiomics with deep-learning, Scientific reports 7 (1) (2017) 1–12.

[11] L. M. Hadjiiski, K. H. Cha, R. H. Cohan, H.-P. Chan, E. M. Caoili, M. S. Davenport, R. K. Samala, A. Z. Weizer, A. Alva, G. Kirova-Nedyalkova, et al., Intraobserver variability in bladder cancer treatment response assessment with and without computerized decision support, Tomography 6 (2) (2020) 194–202.

[12] A. P. Dempster, Upper and lower probabilities generated by a random closed interval, The Annals of Mathematical Statistics (1968) 957–966.

[13] G. Shafer, A mathematical theory of evidence turns 40, International Jour- nal of Approximate Reasoning 79 (2016) 7–25.

[14] S. Antoni, J. Ferlay, I. Soerjomataram, A. Znaor, A. Jemal, F. Bray, Blad- der cancer incidence and mortality: a global overview and recent trends, European urology 71 (1) (2017) 96–108.

[15] D. S. Kaufman, W. U. Shipley, A. S. Feldman, Bladder cancer, The Lancet 374 (9685) (2009) 239–249.

[16] M. J. Magers, A. Lopez-Beltran, R. Montironi, S. R. Williamson, H. Z. Kaimakliotis, L. Cheng, Staging of bladder cancer, Histopathology 74 (1) (2019) 112–134.

[17] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems 25 (2012).

[18] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (7553) (2015) 436–444.

[19] D. Nie, Y. Gao, L. Wang, D. Shen, Asdnet: attention based semi-supervised deep networks for medical image segmentation, in: International conference on medical image computing and computer-assisted intervention (2018) 370–378.

[20] X. Huang, X. Yue, Z. Xu, Y. Chen, Integrating general and specific priors into deep convolutional neural networks for bladder tumor segmentation, in: 2021 International Joint Conference on Neural Networks (2021) 1–8.

[21] K. H. Cha, L. M. Hadjiiski, H.-P. Chan, E. M. Caoili, R. H. Cohan, A. Weizer, R. K. Samala, Computer-aided detection of bladder masses in ct urography (ctu), in: Medical Imaging 2017: Computer-Aided Diagnosis 10134 (2017) 9-13.

[22] E. Shkolyar, X. Jia, T. C. Chang, D. Trivedi, K. E. Mach, M. Q.-H. Meng, L. Xing, J. C. Liao, Augmented bladder tumor detection using deep learning, European urology 76 (6) (2019) 714–718.

[23] F. Rundo, C. Spampinato, G. L. Banna, S. Conoci, Advanced deep learning embedded motion radiomics pipeline for predicting anti-pd-1/pd-l1 immunotherapy response in the treatment of bladder cancer: preliminary results, Electronics 8 (10) (2019) 1134.

[24] K. Hammouda, F. Khalifa, A. Soliman, M. Ghazal, M. Abou El-Ghar, M. Badawy, H. Darwish, A. Khelifi, A. El-Baz, A multiparametric mri-based cad system for accurate diagnosis of bladder cancer staging, Computerized Medical Imaging and Graphics 90 (2021) 101911.

[25] C. Zhang, X. Yue, Y. Chen, Y. Lv, Integrating diagnosis rules into deep neural networks for bladder cancer staging, in: Proceedings of the 29th ACM International Conference on Information and Knowledge Management (2020) 2301–2304.

[26] D. Liu, S. Wang, J. Wang, The effect of ct high-resolution imaging diagnosis based on deep residual network on the pathology of bladder cancer classification and staging, Computer Methods and Programs in Biomedicine 215 (2022) 106635.

[27] T. Denoeux, A k-nearest neighbor classification rule based on dempster-shafer theory, in: Classic works of the Dempster-Shafer theory of belief functions (2008) 737–760.

[28] B. Quost, T. Denoeux, S. Li, Parametric classification with soft labels using the evidential em algorithm: linear discriminant analysis versus logistic regression, Advances in Data Analysis and Classification 11 (4) (2017) 659–690.

[29] T. Denoeux, Logistic regression, neural networks and dempster–shafer theory: A new perspective, Knowledge-Based Systems 176 (2019) 54–67.

[30] A. Capelle, C. Fernandez-Maloigne, O. Colot, Segmentation of brain tumors by evidence theory: on the use of the conflict information, in: International Conference on Information Fusion (2004) 264–271.

[31] C. Lian, S. Ruan, T. Denoeux, H. Li, P. Vera, Dempster-shafer theory based feature selection with sparse constraint for outcome prediction in cancer therapy, in: International conference on medical image computing and computer-assisted intervention (2015) 695–702.

[32] L. Huang, S. Ruan, T. Denoeux, Belief function-based semi-supervised learning for brain tumor segmentation, in: 2021 IEEE 18th International Symposium on Biomedical Imaging (2021) 160–164.

[33] S. Ben Atitallah, M. Driss, W. Boulila, A. Koubaa, H. Ben Ghézala, Fusion of convolutional neural networks based on dempster–shafer theory for automatic pneumonia detection from chest x-ray images, International Journal of Imaging Systems and Technology 32 (2) (2022) 658–672.

[34] M. Sensoy, L. Kaplan, M. Kandemir, Evidential deep learning to quantify classification uncertainty, Advances in Neural Information Processing Systems 31 (2018).

[35] A. Amini, W. Schwarting, A. Soleimany, D. Rus, Deep evidential regression, Advances in Neural Information Processing Systems 33 (2020) 14927–14937.

[36] X. Yue, Y. Chen, B. Yuan, Y. Lv, Three-way image classification with evidential deep convolutional neural networks, Cognitive Computation (2021) 1–13.

[37] B. Yuan, X. Yue, Y. Lv, T. Denoeux, Evidential deep neural networks for uncertain data classification, in: International Conference on Knowledge Science, Engineering and Management (2020) 427–437.

[38] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, (2016) 770–778.

[39] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition (2017) 4700–4708.

[40] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-cam: Visual explanations from deep networks via gradient-based localization, in: Proceedings of the IEEE international conference on computer vision (2017) 618–626.