University of Dundee

Cross-View Self-Similarity Using Shared Dictionary Learning for Cervical Cancer Staging

Bnouni, Nesrine; Rekik, Islem; Rhim, Mohamed Salah; Amara, Najoua Essoukri Ben

Published in:
IEEE Access

DOI:
10.1109/ACCESS.2019.2902654

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):
Bnouni, N., Rekik, I., Rhim, M. S., & Amara, N. E. B. (2019). Cross-View Self-Similarity Using Shared Dictionary Learning for Cervical Cancer Staging. IEEE Access. https://doi.org/10.1109/ACCESS.2019.2902654

Cross-View Self-Similarity Using Shared Dictionary Learning for Cervical Cancer Staging

Nesrine Bnouni, Islem Rekik, *Member, IEEE,* Mohamed Salah Rhim, and Najoua Essoukri Ben Amara

Abstract—Dictionary Learning (DL) has gained large popularity in solving different computer vision and medical image problems. However, to the best of our knowledge, it has not been used for cervical tumor staging. More importantly, there have been very limited works on how to aggregate different *interactions across data views* using dictionary learning. As contribution, we propose a novel Cross-View Self-Similarity Low Rank Shared Dictionary Learning -based (CVSS-LRSDL) framework, which introduces three major contributions in medical image-based cervical cancer staging: (1) leveraging the complementary of axial and sagittal T2w-Magnetic Resonance (MR) views for cervical cancer diagnosis, (2) introducing Self-Similarity (SS) patches for DL training, which explore the *unidirectional interaction* from a source view to a target one, and (3) extracting features that are *shared across* tumor grades as well as *grade-specific* features using the CVSS-LRSDL learning approach. For the first and second contributions, given an input patch in the source view (axial T2w-MR images), we generate its SS patches within a fixed neighborhood in the target view (sagittal T2w-MR images). Specifically, we produce a unidirectional patch-wise SS from a source to a target view, based on mutual and additional information between both views. As for the third contribution, we represent each individual subject using the weighted distance matrix between views, which is used to train our dictionary learning-based classifier to output the label for a new testing subject. Overall, our framework outperformed several DL based multi-label classification methods trained using: (i) patch intensities, (ii) SS single-view patches, and (iii) weighted-SS single-view patches. We evaluated our CVSS-LRSDL framework using 15 T2w-MRI sequences with axial and sagittal views. Our CVSS-LRSDL significantly (p<0.05) outperformed several comparison methods and obtained an average accuracy of 81.73% for cervical cancer staging.

Index Terms—Shared dictionary learning, low-rank models, self-similarity, cross-view, cervical cancer stage.

I. INTRODUCTION

CERVICAL cancer is a recurrent gynecological malignancy and a communal cause of death. The patient outcome relates to the tumor stage, size, nodal status, and histological grade. Staging in cervical cancer highly influences the patient treatment as well as prognosis. The International Federation of Gynecology and Obstetrics (FIGO) [1] staging system is the most used for cervical cancer. Knowing the stage helps the doctor to decide what kind of treatment is best to help predict a patient's prognosis [2], [3]. Cervical cancer staging is a complex task as it covers different grades [4]. Generally, high grades are correlated with the extent of cancer spread. Magnetic Resonance Imaging (MRI) is a non-invasive imaging technique that is commonly used for tumor detection, staging and measurement of its exact volume, shape, and local extent of the disease, thereby assisting clinicians in treatment planning. In particular, T2-weighted (T2w) MRI images play a crucial role in the identification of the primary tumor and assessment of its extent as it provides accurate anatomic details and a fine contrast resolution. Exploiting axial-T2w and sagittal T2w views is adequate for staging in most cases. T2w-MRI images of the pelvis are acquired in axial, sagittal and coronal planes. In cervical cancer, while axial-T2w images arranged perpendicular to the long axis of cervix yield more precise evaluation of the parametrial invasion, the nodes status [5] and the stromal involvement, sagittal-T2w images planned parallel to the long axis of cervix provide more accurate assessment of the tumor size [6], [7] and the extension of neighboring organs.

FIGO divides cervical cancer into four stages. Each stage is further divided into several sub-stages. Specifically, stage I tumor is restricted to the cervix. Stage IA is a microscopic disease and is not observable on MRI. An observable tumor, even with superficial invasion, is stage IB cancer. Stage II is well-inspected when the tumor spreads over the cervix. Involvement of the upper two thirds of the vagina identify IIA stage. According to FIGO staging, if the tumor area is $\leq$ 4 cm, the patient is staged as IIA1, otherwise the patient is staged as IIA2. In stage IIB, the tumor perturbs the normally hypointense peripheral stroma on T2w-MR images and spreads in the parametrium. Stage III is established when the tumor extends to the lower third of the vagina or the lateral pelvic wall with related hydronephrosis. The bladder and/or rectal mucosa involvement or distant metastasis specifies stage IV. To the best of our knowledge, there has been no record in the cervical cancer literature of the use or development of automatic classification methods for cervical cancer staging.

On the other hand, Sparse Dictionary Learning (SDL) strategies have been developed to solve multi-label classification problems by learning how to extract relevant information. They have focused on training data from various classes into a unique dictionary and uses class-specific residue for classification. This has been achieved by forcing some sort of sparseness constraints on the coefficients of the learned representation [8]. Specifically, SDL is a particular sparse

N. Bnouni is affiliated with LATIS, Laboratory of Advanced Technology and Intelligent Systems, ENISo, Sousse University, Tunisia and co-affiliated with BASIRA Lab, Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey.

I. Rekik is affiliated with BASIRA Lab, Faculty of Computer and Informatics Engineering, Istanbul Technical University, Istanbul, Turkey and co-affiliated with the School of Science and Engineering, University of Dundee, Dundee, UK.

M.S Rhim is affiliated with the Department of Gynecology Obstetrics, Faculty of Medicine of Monastir, Tunisia.

N. Essouki Ben Amara is affiliated with LATIS, Laboratory of Advanced Technology and Intelligent Systems, ENISo, Sousse University, Tunisia.

signal representation and has increased in recent years. Its goal is to learn a dictionary in which only a few atoms can be linearly combined to represent a given signal. SDL-based approaches have performed state-of-the-art performances in many applications including aperture radar image classification [9], sparse signal recovery [10], image segmentation [11], image de-noising [12], video anomaly detection [13] recaptured image recognition [14], hyper-spectral image classification [15], [16], [17] and medical image classification [18], [19], [20]. An SDL representation-based classifier has been also applied for robust face recognition [21]. However, the majority of existing classification-based SDL approaches have aimed to learn discriminative class-specific dictionaries by forcing block-diagonal constraints or favoring the incoherence between class-specific dictionaries [22], [23], [20], [24], [8], [25], [26], [27], [19]. This is a strong constraint that overlooks the shared features across classes. This problem has been partially addressed by recent efforts [28], [29], [30], [31]. However, shared features across classes have been masked in the learned class-specific dictionaries. To address this issue, Low-Rank Shared Dictionary Learning (LRSDL) framework was proposed in [32], [33] with the special capability of capturing class-specific features as well as shared features across different classes. It was essentially a generalized variant of Fisher Discrimination Dictionary Learning (FDDL)[34], [23] with the capacity of learning shared features.

In this paper, we first aim to adapt LRSDL to automatically learn stage-specific dictionaries and a shared dictionary across cervical cancer stages. Specifically, each stage-dictionary per stage is simultaneously learned to extract discriminative features as well as the shared features that all cancer stages contain. For the shared part, two intuitive constraints are imposed. First, the shared dictionary must contain discriminative grade features due to the low-rank structure. Second, the sparse coefficients corresponding to the shared dictionary must be similar. Both of these constraints are capital for the shared dictionary. On the other hand, patch-based image modeling has gained a wide popularity in computer vision and medical images. In particular, the use of image Self-Similarity (SS) which implies that a patch often has many similar patches across the image, has significantly improved the performance many applications. Specifically, SS is used to approximate each patch, in some way, by other patches and to indicate how much exactly or approximately patches are similar to a part of them to evaluate the pairwise relationship. Particularly, the ability to compare patches has been the basis of many approaches in computer vision problems. Patches in medical images are statistically self-similar. For instance, in [35], learning similarity between cross-spectral patches with a 2-channel convolutional neural network model was proposed. In [36], a global SS and its advantages over local SS was explored. In [37], a patch group-based non-local SS prior-learning scheme to learn explicit nonlocal SS models from natural images for high-performance de-noising was proposed. Crus et al. [38] introduced a SS based approach which used large groups of similar patches extracted from the input image for image super-resolution. Bustin et al. [39] put forward a novel isotropic 3D reconstruction scheme from brain imaging

and clinical cardiac MRI that integrated non-local and SS information from 3D patch neighborhoods. Manjon et al. [40] proposed a new super-resolution method to reconstruct high-resolution MR images from the low-resolution images using SS and image priors. However, these approaches have only explored within-modality (or within-view) similarity, where similar patches are identified within the same view, which overlooks patches similarities across views. In other words, these overlooked inter-view interactions have been shared, and complementary information between views can be leveraged to further boost the target results.

Inspired by such works and to fully exploit SS across views, we propose a novel Cross-view Self-Similarity LRSDL-based (CVSS- LRSDL) method, which leverages the complementary between the source view (axial-MRI) and the target view (sagittal-MRI) for cervical cancer staging using LRSDL-based dictionary learning trained on SS cross-view patches. Specifically, given a patch in the source view (axial T2w-MRI), we extract its SS patches within a fixed neighborhood in the target view (sagittal T2w-MRI), thereby defining a *unidirectional cross-view* information transfer. Since the sagittal T2-w image is more reliable to stage cervical cancer and contains more spatial information such as tumor size and extension to neighborhood organs, we fixed it as the target view where more relevant information is extracted via locally retrieving many similarity patches. The axial T2-w image (source view) remains the basic MRI view for the detection of parametrial invasion in cervical cancer. This disentangles the features that are shared across tumor grades as well as grade-specific features using LRSDL learning approach. CVSS- LRSDL can jointly learn the features from cross-view by transferring sparse feature representations of patches from the source to the target view to improve cervical cancer staging. Our method has two major contributions to the state-of-the-art of cervical cancer staging as well as DL for multi-label classification:

1. CVSS- LRSDL method learns a shared dictionary across grades as well as grade-specific dictionaries, which allows explicitly and simultaneously learning a set of common cervical cancer staging patterns (the tumor size, how far the tumor went into invading surrounding tissues in and out of the cervix, and its propagation to distant organs) in addition to identifying distinct grade-specific features to determine the cervical cancer stage.

2. CVSS- LRSDL introduces a unidirectional patch-wise SS from a source towards a target view, which allows capturing shared and complementary information between different views. The intuition behind selecting a unidirectional SS lies in the assumption that the target view holds more relevant information for the target classification task compared with the source view because for a single source patch many neighboring self-similar patches are extracted from the target view. Prior to detailing our framework, we give an overview of landmark dictionary learning approaches for multi-label classification.

A- Cross-view Self-Similarity CVSS
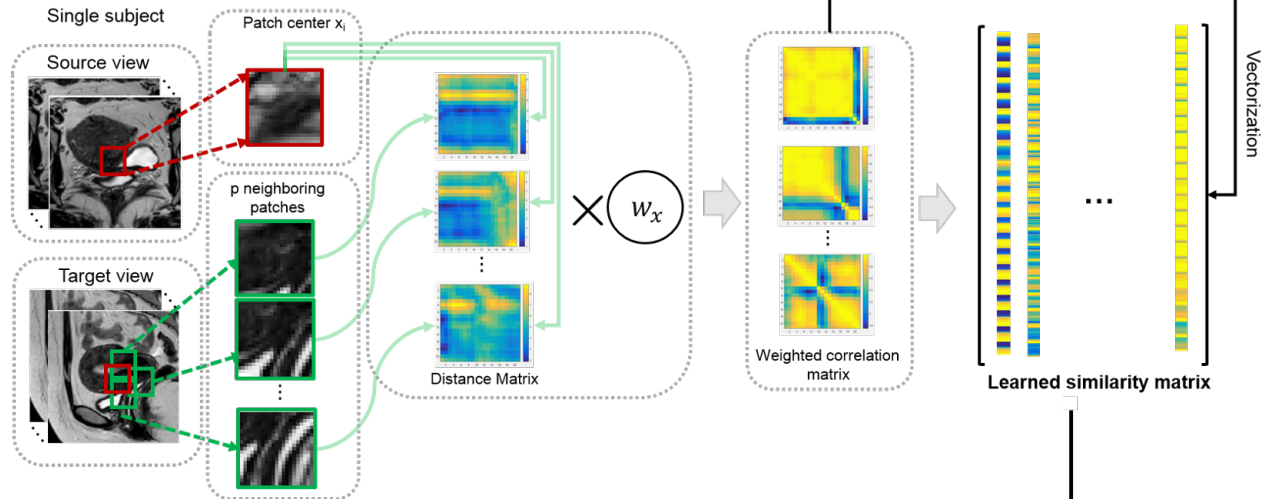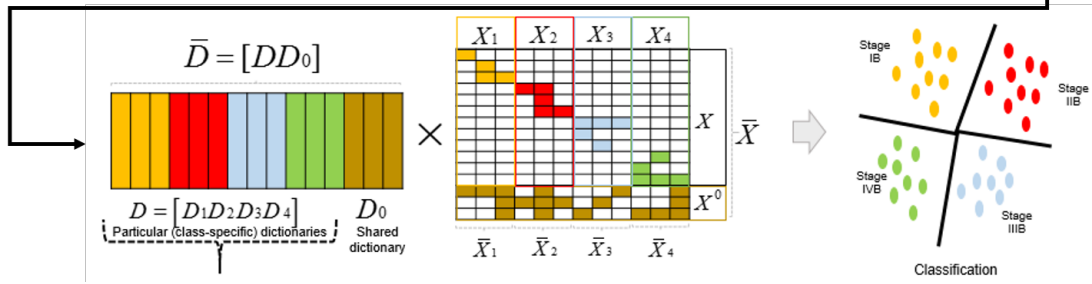

B- LRSDL for classification

Fig. 1: *Pipeline of proposed CVSS-LRSDL using both axial (source) and sagittal (target) views for cervical cancer staging. A- CVSS- LRSDL presents a unidirectional patch-wise SS from a source (axial-T2w) towards a target view (sagittal-T2w), allowing to better capture shared and complementary information between both views. Specifically, we extract the source view and the target view sagittal patches. Then, for each input patch in the source view (axial T2w-MRI), we produce its SS patches within a fixed neighboring patches in the target view (sagittal T2w-MRI) using a weighted exponential decay of the Minkowski distance. B- We use the LRSDL approach on our training samples and their corresponding grades. Next, we test it on a new patient to predict the cancer grade. Specifically, the LRSDL approach considers a shared dictionary across grades and grade-specific dictionaries enabling to explicitly and concurrently learn features of common cervical cancer staging as well as knowing different grade-specific patterns to predict the target cervical cancer grade.*

II. DICTIONARY LEARNING

A. Sparse dictionary learning

SDL has gained popularity in learning parsimonious representations of data. It is an essential problem in signal processing and machine learning. The key idea of sparse representation methods is to exhibit a test pattern as a linear combination of samples from the training set. Specifically, sparsity is revealed as most of non-zeros answer to bases, whose memberships are identical to the test sample. Consequently, in the ideal case, each target is anticipated to belong in its proper class subspace, and all class sub-spaces are nonoverlapping. Given a G cancer grade and a dictionary learning $\mathbf{D} = [\mathbf{D}_1, \cdots, \mathbf{D}_G]$ with $\mathbf{D}_g$ comprising training samples from cancer grade g, $g = 1, \cdots, G$, a sample $\mathbf{Y}$ from cancer grade g can be represented as $\mathbf{Y} \approx \mathbf{D}_g \mathbf{X}^g$. Therefore, $\mathbf{Y} \approx \mathbf{D}\mathbf{X} = \mathbf{D}_1 \mathbf{X}^1 + \cdots + \mathbf{D}_g \mathbf{X}^g + \cdots + \mathbf{D}_G \mathbf{X}^G$, then most of active coefficients of $\mathbf{X}$ should be placed in $\mathbf{X}^g$ and the element vector $\mathbf{X}$ is expected to be sparse. In a matrix form,

$\mathbf{Y} = [\mathbf{Y}_1, \cdots, \mathbf{Y}_g, \cdots, \mathbf{Y}_G]$ is the set of samples where $\mathbf{Y}_g$ includes those in cancer grade g, and the element matrix $\mathbf{X}$ is sparse. In the ideal case, $\mathbf{X}$ is a block diagonal structure.

In addition, $\mathbf{D}_0$ is a shared dictionary and $\bar{\mathbf{D}} = [\mathbf{D} \ \mathbf{D}_0]$ is a total dictionary, $\bar{\mathbf{X}} = [\mathbf{X}^{\mathbf{T}}, (\mathbf{X}^0)^T]^T$ and $\bar{\mathbf{X}}_{\mathbf{g}} = [(\mathbf{X}_g)^T, (\mathbf{X}_g^0)^T]^T$. $\mathbf{m}, \mathbf{m}_0, \mathbf{m}_c$ are the mean vectors of $\mathbf{X}, \mathbf{X}_0, \mathbf{X}_g$, respectively. $\mathbf{M}, \mathbf{M}_0, \mathbf{M}_c$ are the mean matrices of $\mathbf{X}, \mathbf{X}_0, \mathbf{X}_g$, respectively.

B. Fisher discrimination dictionary learning

FDDL [34], [23] has been widely used as a method for learning both dictionary structure and discriminative coefficients. Especially, the discriminative Dictionary $\mathbf{D}$ and the sparse coefficient matrix $\mathbf{X}$ are trained based on minimizing the subsequent cost function as follows **(1)**.

$$J(\mathbf{D}, \mathbf{X}) = \frac{1}{2} \sum_{g=1}^{G} r(\mathbf{Y}_g, \mathbf{D}, \mathbf{X_g}) + \lambda_1 \|\mathbf{X}\|_1 + \frac{\lambda_2}{2} f(\mathbf{X}), \quad (1)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2902654, IEEE Access

IEEE ACCESS 2019   4

where $r(\mathbf{Y}_g, \mathbf{D}, \mathbf{X}_g) = \|\mathbf{Y}_g - \mathbf{D}\mathbf{X}_g\|_F^2 + \|\mathbf{Y}_g - \mathbf{D}_g\mathbf{X}_g^g\|_F^2 + \sum_{i \neq g} \|\mathbf{D}_i\mathbf{X}_g^i\|_F^2$,

$\sum_{g=1}^{G} r(\mathbf{Y}_g, \mathbf{D}, \mathbf{X}_g)$ is the discriminative fidelity term,

$f(\mathbf{X}) = \sum_{g=1}^{G} (\|\mathbf{X}_g - \mathbf{M}_g\|_F^2 - \|\mathbf{M}_g - \mathbf{M}\|_F^2) + \|\mathbf{X}\|_F^2$ is the

Fisher-based discriminative coefficient, (λ_1, λ_2) are positive regularization parameters and the l_1-norm promises the sparsity. The minimization of Equation (1) is achieved otherwise by optimizing each variable $\mathbf{X}_g$ or $\mathbf{D}_g$. This technique has a heavy convergence process, which might not be feasible for solving multi-class big-size problems.

C. Separating particularity and commonality dictionary learning

Even with different classes detaining separate class-specific features, classes usually share common patterns. This aspect was clearly overlooked by previous SDL and FDDL methods. To address this limitation, [30] proposed a dictionary learning framework by separating the particularity and the commonality (COPAR). This algorithm can be favorable for optimizing both sub-problems of updating the common dictionary as well as the class-specific dictionaries. COPAR is an extended variant of SDL technique, which considers the shared dictionary without using the low-rank constraint with the following cost function defined in Equation (2).

$$\frac{1}{2}f_1(\mathbf{Y}, \bar{\mathbf{D}}, \bar{\mathbf{X}}) + \lambda\|\bar{\mathbf{X}}\|_1 + \eta \sum_{g=0}^{G} \sum_{\substack{i \neq g \\ i=0}}^{G} \|\mathbf{D}_i^T\mathbf{D}_g\|_F^2, \quad (2)$$

with $f_1(\mathbf{Y}, \bar{\mathbf{D}}, \bar{\mathbf{X}}) = \sum_{g=1}^{G} r_1(\mathbf{Y}, \bar{\mathbf{D}}, \bar{\mathbf{X}}_g), r_1(\mathbf{Y}, \bar{\mathbf{D}}, \bar{\mathbf{X}}_g) =$

$\|\mathbf{Y}_g - \mathbf{D}\mathbf{X}_g\|_F^2 + \|\mathbf{Y}_g - \mathbf{D}_0\mathbf{X}_g^0 - \mathbf{D}_g\mathbf{X}_g^G\|_F^2) + \sum_{\substack{j \neq g \\ j=1}}^{G} \|\mathbf{X}_g^j\|_F^2$

and λ_1 is a positive regularization parameter.

D. Low-rank shared dictionary learning

Learning step. LRSDL [32], [33] is basically a generalized FDDL method with the additional capacity to capture the shared features across classes. It learns both shared and class specific dictionaries. Particularly, a low-rank constraint is imposed that the coefficients of the shared dictionary should be similar and its sub-space should take low dimension. It is anticipated that $\mathbf{Y}_g$, with the presence of the shared dictionary, can be greatly explained by the collaboration between the particular dictionary $\mathbf{D}_g$ and the shared dictionary $\mathbf{D}_0$. Specifically, the discriminative fidelity $r(\mathbf{Y}_g, \mathbf{D}, \mathbf{X}_g)$ in (1) can be extended to $\bar{r}(\mathbf{Y}_g, \bar{\mathbf{D}}, \bar{\mathbf{X}}_g)$ defined in Equation (3).

$$\|\mathbf{Y}_c - \bar{\mathbf{D}}\bar{\mathbf{X}}_g\|_F^2 + \|\mathbf{Y}_g - \mathbf{D}_g\mathbf{X}_g^g - \mathbf{D}_0\mathbf{X}_g^0\|_F^2 + \sum_{\substack{i \neq g \\ i=1}} \|\mathbf{D}_i\mathbf{X}_g^i\|_F^2.$$

$$(3)$$

The Fisher-based discriminative term f(X) is extended to $\bar{f}(\bar{\mathbf{X}})$ outlined as follows (4).

$$\bar{f}(\bar{\mathbf{X}}) = f(X) + \|\mathbf{X}^0 - \mathbf{M}^0\|_F^2, \quad (4)$$

where $\|\mathbf{X}^0 - \mathbf{M}^0\|_F^2$ constrains the coefficients of the training samples represented by the shared dictionary to be equal. For the shared dictionary, rank $(\mathbf{D}_0)$ is constrained to be small by using the low-rank constraint using the nuclear norm $\|\mathbf{D}_0\|_*$. Ultimately, the cost function $\bar{J}(\bar{\mathbf{D}}, \bar{\mathbf{X}})$ of the LRSDL is as in Equation (5).

$$\frac{1}{2}\sum_{g=1}^{G} \bar{r}(\mathbf{y}_g, \bar{\mathbf{D}}, \bar{\mathbf{X}}_g) + \lambda_1\|\bar{\mathbf{X}}\|_1 + \frac{\lambda_2}{2}\bar{f}(\bar{\mathbf{X}}) + \eta\|\mathbf{D}^0\|_*, \quad (5)$$

where $(\lambda_1, \lambda_2, \eta)$ are positive regularization parameters. By minimizing this unbiased function, the appropriate dictionaries can be found. Particularly, if $k_0 = 0$, then $\bar{\mathbf{D}}, \bar{\mathbf{X}}$ becomes $\mathbf{D}, \mathbf{X}$, and respectively $\bar{J}(\bar{\mathbf{D}}, \bar{\mathbf{X}})$ becomes $J(\mathbf{D}, \mathbf{X})$ and the LRSDL becomes FDDL.

Classification step. As a result of the learning process, we estimate $\bar{\mathbf{D}}, \mathbf{m}_g, \mathbf{m}^0$. For a new test sample $\mathbf{y}$, we first locate its coefficient vector $\bar{\mathbf{x}}$ and constrain $\mathbf{x}^0$ to be close to $\mathbf{m}^0$ as in Equation (6).

$$\bar{\mathbf{x}} = \underset{\bar{\mathbf{x}}}{\arg\min} \frac{1}{2}\|\mathbf{y} - \bar{\mathbf{D}}\bar{\mathbf{x}}\|_2^2 + \frac{\lambda_2}{2}\|\mathbf{x}^0 - \mathbf{m}^0\|_2^2 + \lambda_1\|\bar{\mathbf{x}}\|_1. \quad (6)$$

Then, we extract the shared dictionary to obtain $\bar{\mathbf{y}} = \mathbf{y} - \mathbf{D}_0\mathbf{x}^0$. $\mathbf{y}$ is estimated by solving Equation (7).

$$\underset{1 \leq g \leq G}{\arg\min}(w\|\bar{\mathbf{y}} - \mathbf{D}_g\mathbf{x}^g\|_2^2 + (1-w)\|\mathbf{x} - \mathbf{m}_g\|_2^2), \quad (7)$$

where $w \in [0, 1]$ is a preset weight for balancing the contribution of the two terms.

III. PROPOSED CROSS-VIEW SELF-SIMILARITY BASED DICTIONARY LEARNING FOR TUMOR STAGING

In this section, we detail the steps of our novel CVSS-LRSDL-based framework. **Fig. 1** illustrates the different steps of the proposed classification framework. First, we extract the source view (axial T2w-MRI) and the target one (sagittal T2w-MRI) patches of cervical MR images. Second, for an input patch in the source view, we generate its SS patches within a fixed neighborhood in the target view using a weighted exponential decay of the Minkowski distance. Finally, we train our CVSS-LRSDL on training samples and their corresponding grades, then test it on a new patient to predict the cancer grade. We detail in the following the various steps of the proposed CVSS-LRSDL multi-label classification framework. Specifically, we use two MRI views including the source view (axial-T2w) and the target view (sagittal-T2w). Next, we look for spatial neighbors not in the current view but in the target view, building a *unidirectional cross-view information* transfer. Particularly, for each location $\mathbf{x}$ in the source view, we extract a patch $\mathbf{p}_s(\mathbf{x})$ of size $n \times n$. Each patch $\mathbf{p}_s(\mathbf{x})$ centered at $\mathbf{x}$.

Next, we extract k neighboring patches $\mathbf{p}_t(\mathbf{x}_i)$ $(i = 1, 2 \cdots k)$ around location $\mathbf{x}$ in the target view.

Then we compute at each location $\mathbf{x}$ the similarity patch between $\mathbf{p}_s(\mathbf{x})$ in the source view and its k nearest neighboring intensity patches $\mathbf{p}_t(\mathbf{x}_i)$ $(i = 1, 2 \cdots k)$ in the target view using the Minkowski distance $\mathbf{M}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$, respectively. We define the patch-based Minkowski distance as the sum of the p-order pairwise differences between patches $\mathbf{p}_s(\mathbf{x})$ and $\mathbf{p}_t(\mathbf{x}_i)$ exponentiated to the power of $\frac{1}{p}$ as in Equation **(8)**.

$$\mathbf{M}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) = (\sum_{j=1}^{n} |\mathbf{p}_s(\mathbf{x})(j) - \mathbf{p}_t(\mathbf{x}_i)(j)|^p)^{\frac{1}{p}}. \quad (8)$$

Finally, at each location $\mathbf{x}$, we propose a weighted SS using weighted correlation coefficient between each similarity patch $\mathbf{M}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ and a pixel-wise weight vector $\mathbf{w}_x$ of size $n \times 1$ [41]. Specifically, the predefined similarity patch is calculated using the element-wise weighted distance between the source patch located at $\mathbf{x}$ and its k corresponding neighboring patches. Exponential decay is used to yield more reliable and robust dynamic measures to reduce the autocorrelation of dynamic correlations while keeping significance and robustness of the measure [41]. This is the type of weighting we use in this work, however other kinds can be considered as well. Weights are therefore defined as follows **(9)**.

$$\mathbf{w}_\mathbf{x}(j) = w_0 e^{\alpha(j-n)}, \forall j \in \{1, 2, ..., n\} \quad (9)$$

where $\alpha \in \mathbb{R}$, $\alpha \geq 0$ denotes is the exponential decay factor with Equation **(10)**:

$$\alpha = 2/n, \quad (10)$$

The weighted correlation coefficients satisfy the following criteria **(11)**.

$$\mathbf{w}_\mathbf{x}(j) \geq 0, \sum_{j=1}^{n} \mathbf{w}_\mathbf{x}(j) = 1 \quad (11)$$

The given w_0 in Equation **(9)** can be set by observing that $\mathbf{w}$ is subject to the constraint outlined in Equation **(11)**.

$$\sum_{t=1}^{n} w_0 e^{\alpha(j-n)} = 1.$$

By solving w_0, $w_0(\alpha) = \frac{1}{\sum_{j=1}^{n} e^{\alpha(j-n)}}$

as $\sum_{j=1}^{n} e^{\alpha(j-n)} = \frac{e^\alpha}{e^\alpha - 1}(1 - e^{-\alpha n})$

In fact:
$$\sum_{j=1}^{n} e^{\alpha(j-n)} = e^{-\alpha n} \sum_{j=1}^{n} (e^\alpha)^j = e^{-\alpha n}(\frac{1-(e^\alpha)^{n+1}}{1-(e^\alpha)} - 1)$$

$$= \frac{e^{-\alpha n}}{1-e^\alpha}(e^\alpha - e^{n+\alpha}) = \frac{e^\alpha}{e^\alpha - 1}(1 - e^{-\alpha n}).$$

Subsequently, $w_0(\alpha) = \frac{1-e^{-\alpha}}{1-e^{-\alpha n}}$. When α is 0 as in Equation **(12)**, the weights are uniform. When α is large as in Equation **(13)**, spatially distant pixels become increasingly irrelevant and neighboring ones become ever more important.

$$\lim_{\alpha \to 0^+} w_0(\alpha) = \frac{1}{n} \quad (12)$$

$$\lim_{\alpha \to +\infty} w_0(\alpha) = 1 \quad (13)$$

After obtaining the weighted coefficients $\mathbf{w}_x$, we compute at each location $\mathbf{x}$ the self-similarity patch $\mathbf{S}_{\mathbf{w}_x}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ between $\mathbf{p}_s(\mathbf{x})$ in the source view and its k nearest neighboring intensity patches $\mathbf{p}_t(\mathbf{x}_i)$ $(i = 1, 2 \cdots k)$ in the target view using Pearson weighted correlation [41] as in the following Equation:

$$\mathbf{S}_{\mathbf{w_x}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) = \frac{\sigma_{\mathbf{w}_x}^{\mathbf{uv}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))}{\sigma_{\mathbf{w}_x}^{\mathbf{u}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))\sigma_{\mathbf{w}_x}^{\mathbf{v}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))} \quad (14)$$

It is calculated using weighted means, weighted variances and weighted covariances:

$$\bar{\mathbf{m}}_{\mathbf{w_x}}^{u}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) = \sum_{j=1}^{n} \mathbf{w}_\mathbf{x}(j)\mathbf{m}^u(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)),$$ where the

distance $\mathbf{m}^u(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ represents the uth column vector of matrix $\mathbf{M}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ and $u \in \{1, \ldots n\}$.

$$\sigma_{\mathbf{w}_x}^{u}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) =$$
$$\sqrt{\sum_{j=1}^{n} \mathbf{w}_\mathbf{x}(j)[\mathbf{m}^u(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) - \bar{\mathbf{m}}_{\mathbf{w_x}}^{u}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))]^2}$$

$$\sigma_{\mathbf{w_x}}^{\mathbf{uv}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) = \sum_{j=1}^{n} \mathbf{w}_\mathbf{x}(j)([\mathbf{m}^u(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$$
$$- \bar{\mathbf{m}}_{\mathbf{w_x}}^{u}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))][\mathbf{m}^{\mathbf{v}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i)) - \bar{\mathbf{m}}_{\mathbf{w_x}}^{\mathbf{v}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))]).$$

where $\sigma_{\mathbf{w_x}}^{\mathbf{uv}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ is the covariance of the two variables $\mathbf{m}^{\mathbf{u}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$ and $\mathbf{m}^{\mathbf{v}}(\mathbf{p}_s(\mathbf{x}), \mathbf{p}_t(\mathbf{x}_i))$.

Finally, we vectorize all cross-view self-similarity patches from all training subjects and combine them in column vectors into a similarity matrix to train the dictionary learning and obtain the final classification result. In practice, given G cancer grade classes, $g = \{1, \cdots, G\}$ and dictionary $\mathbf{D} = [\mathbf{D}_1, \cdots, \mathbf{D}_G]$, each subject is located in its own subspace, and all class sub-spaces, in the perfect case, are non-overlapping. Our proposed CVSS-LRSDL classification framework integrates a unidirectional similarity from a source to a target view to enhance cervical cancer detection and further reduce the tumor grading errors. Specifically, our CVSS-LRSDL has been trained using all CVSS training samples and their corresponding stages. In the testing stage, each CVSS testing sample extracted from the patches (source view) and their corresponding neighboring patches (target view) passes through our trained LRSDL to predict the cancer grade **(Fig. 1)**.

IV. RESULTS & DISCUSSION

A. Evaluation Data and Comparison Methods

Dataset. We evaluated the proposed classification framework using 15 clinical pelvic T2-w MR database with different stages (3 case stage IB, 5 case stage IIB, 4 case stage IIIB and 3 case stage IVA) **(Fig. 2)**, acquired between January 2016 and December 2017 diagnosed with cervical cancer. All 15 patients underwent a regular examination with a 1.5-T MRI system. Axial and sagittal T2w 2D turbo spin-echo sequences with voxel size of 0.5×0.5×3 mm were used. All data sets were analyzed and staged by an expert radiologist.
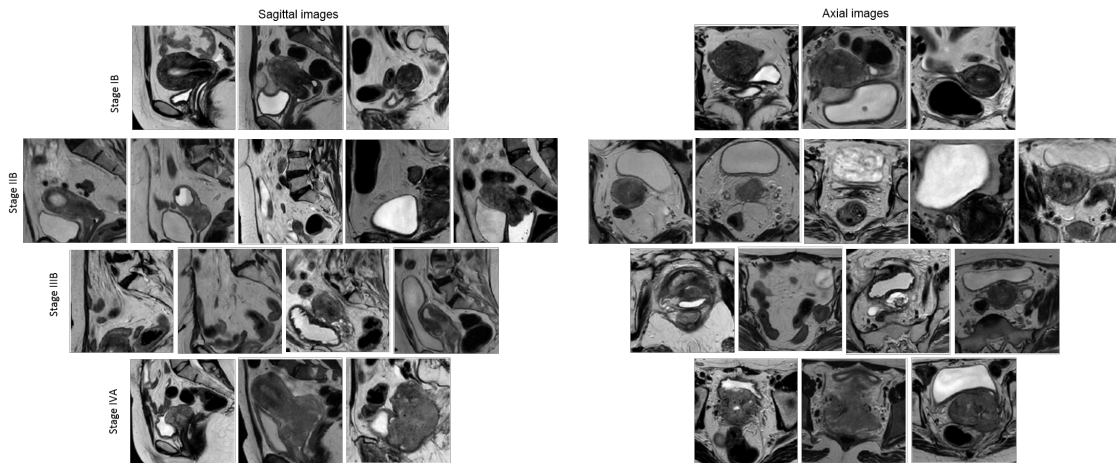
Fig. 2: Axial and sagittal T2w views in 15 pelvic MR images with different cancer grades.

We used affine registration to align all training MR images to a common space. Each testing subject is then affinely aligned to the training samples. To speed up DL training, we extracted intensity patches within a bounding box of size 120×120 automatically drawn around the cervix and its neighboring regions on axial and sagittal views.

Comparison methods. To demonstrate the effectiveness of integrating CVSS into a single framework, we benchmarked our method against approaches without CVSS and when only using SS. Specifically, for evaluation, four types of features were used to train three dictionary learning strategies using: (1) the proposed cross-view self-similarity (CVSS), (2) the weighted SS within the sagittal view (i.e., without cross-view similarity), (3) SS without weighting and computed within the sagittal view, and (4) raw intensity patches in the sagittal view.

B. Experiment setup

We used leave-one-out cross-validation to evaluate our proposed method and its comparison methods. Specifically, for each training subject, we extracted 2D intensity patches of size of 21×21 and 23×23 from each axial and sagittal view for each training subject. For patches of size 21×21 centered at location $\mathbf{x}$, we computed a cross-view SS matrix of size 441×81. Particularly, we extracted $k = 81$ neighboring patches for each patch center and $441 = 21 \times 21$ features for each patch. For patches of size 23×23, we extracted a learned similarity matrix of size 529×100. The classification accuracy was evaluated using average accuracy (%) computed as follows (**15**):

$$ACC = \frac{TP + TN}{TP + FN + TN + FP} \qquad (15)$$

Where TP, TN, FP and FN denotes the number of true positives, true negatives, false positives and false negatives, respectively. The performance of classification results generated by different dictionary learning methods was assessed using the ground truth stage label value (1, 2, 3, 4), (i.e., stage IB, stage IIB, stage IIIB, stage IVA). The classification stability was evaluated by reporting the average accuracy of 50 runs of our algorithm.

C. Performance of proposed method

Fig. 3 displays our learned shared $\mathbf{D_0}$ and class-specific dictionaries $\mathbf{D}$. Samples of different grade of cervical cancer are shown in **Fig. 3**. This figure shows sample learned bases using LRSDL where the shared bases are extracted and collected in the shared dictionary. Class-specific elements and shared elements are nicely decomposed into appropriate sub-dictionaries due to the low-rank constraint $\mathbf{X_0}$ on the shared dictionary of LRSDL imposing thereafter, the block-diagonal constraints on $\mathbf{X}$. The classification results are listed in **Table I**. We observe that the performance of only using intensity patches in almost all methods is boosted when integrating the Minkowski patch difference. We also note that the weight improves classification results, while the use of CVSS boosts the performance going up to 7%. This demonstrates the effectiveness of the proposed weighted CVSS-DL in improving the performance of baseline DL approaches. **Fig. 4** and **Fig. 5** show the overall classification results of different DL methods evaluated on the cervical cancer dataset in terms of the classification accuracy for patch size equals to 21×21 and 23×23 , respectively. Three dictionary learning methods with shared features are used. **Table I**, **Fig. 4** and **Fig. 5** compare our proposed CVSS with several state-of-the-art dictionary learning algorithms including: CVSS-FDDL, CVSS-COPAR and CVSS-LRSDL. Overall, all results derived from comparison methods perform reasonably well on predicting the stage of the cervical cancer. Our proposed CVSS-LRSDL framework significantly outperformed all comparison methods (p<0.05 using paired two tailed t-test) and generated the best classification accuracy. We note a significant increase in the classification accuracy from 47.33% using the intensity patch up to 81.73 % when using the SS using weighted correlation distance.

D. Sensitivity of parameters

Fig. 6 displays the classification accuracy of three dictionary learning classification methods (FDDL, COPAR and LRSDL) using pixel intensity while varying the size of the patch. **Fig. 6**
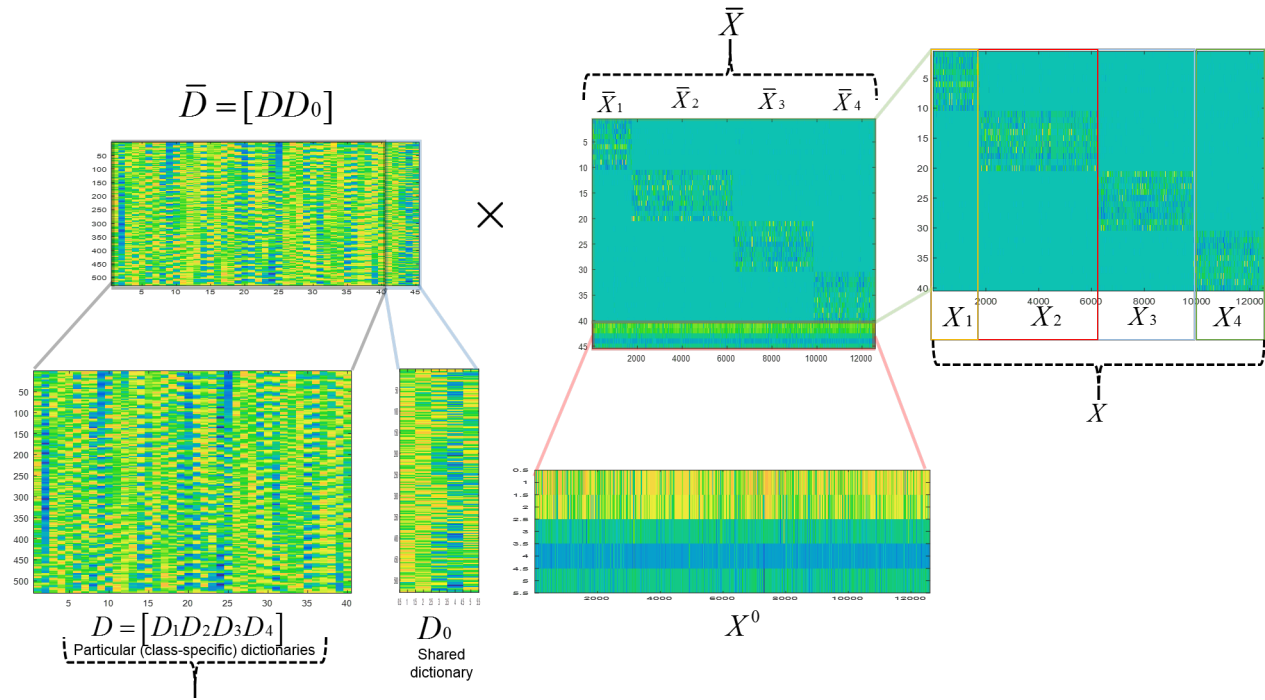
Fig. 3: Visualization of learned shared and grade-specific dictionaries.

TABLE I: Overall accuracy with patch size of 21×21 and 23×23, respectively. Accuracy of different dictionary learning methods for cervical cancer grading.

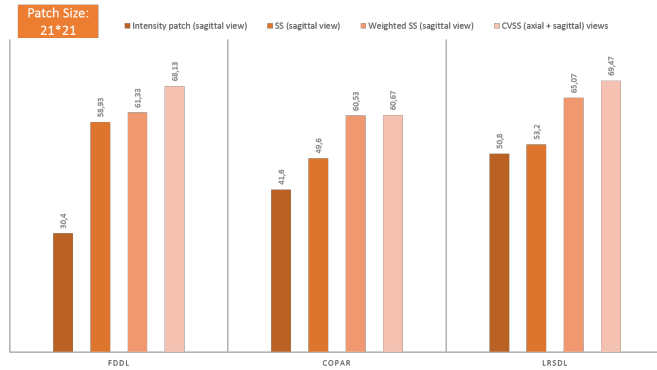| Patch size	Method	Intensity patch (sagittal view)(%)	SS (sagittal view)(%)	Weighted SS (sagittal view)(%)	CVSS (axial + sagittal) views (%)
21*21	FDDL	30.4	58.93	61.33	68.13
21*21	COPAR	41.6	49.6	60.53	60.67
21*21	LRSDL	50.8	53.2	65.07	**69.47**
23*23	FDDL	29.47	77.47	77.6	77.6
23*23	COPAR	64.93	62.27	67.33	68
23*23	LRSDL	47.33	70.93	81.60	**81.73**

Fig. 4: Classification accuracy using 21×21 patch size. FDDL, COPAR and LRSDL are trained using different measures including (i) raw intensity patches in the sagittal view, (ii) SS without weighting and computed within the sagittal view, (iii) the weighted SS within the sagittal view, and (iv) the proposed cross-view self-similarity (CVSS).
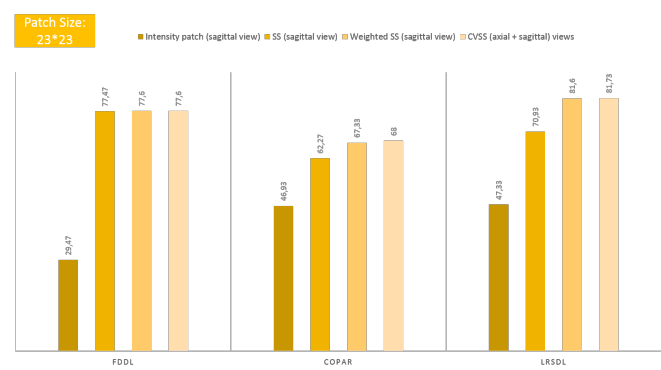

Fig. 5: Classification accuracy using 23×23 patch size. FDDL, COPAR and LRSDL are trained using different measures including (i) raw intensity patches in the sagittal view, (ii) SS without weighting and computed within the sagittal view, (iii) the weighted SS within the sagittal view, and (iv) the proposed cross-view self-similarity (CVSS)
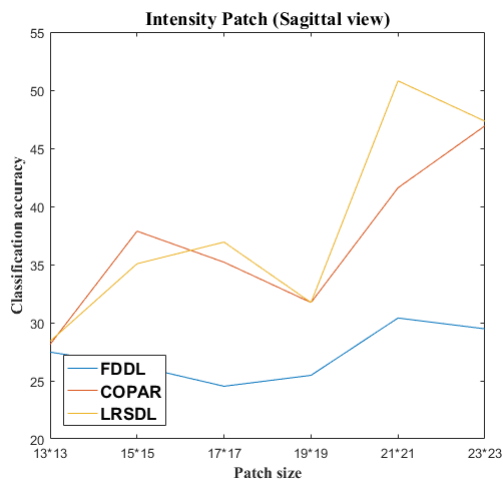
Fig. 6: Classification results with different size of intensity patch. This graph displays the classification results by three dictionary learning approaches using five different patch sizes ranging between 15×15 and 23×23 in the sagittal view.

shows that larger patches produce better results. Thus, we set the patch size to 21×21 and 23×23.

E. Limitations & future work

Our proposed CVSS- LRSDL framework for cervical cancer grading has essentially two strengths: (1) a set of common cervical cancer staging patterns was learned and distinct grade-specific features due to a shared dictionary across grades as well as grade-specific dictionaries to decide cervical cancer stage, and (2) a unidirectional patch-wise SS from a source towards a target view permits to catch joint and additional information among different views. However, our method has a few limitations. First, the multi-scale aspect of the tumor lesions is overlooked in our framework since the patch size is fixed. To address this limitation, we will further extend CVSS-LRSDL by learning scale-specific dictionaries that capture tumor details at different scales for improving the classification accuracy. Second, so far, we have only considered a unidirectional interaction from a source view to a target view given that the target view holds more relevant information for the classification in hand. To generalize our framework in the case where the views hold equally relevant information that is complementary, we can define a bidirectional flow across views in the spirit of [42], [6] where different classifiers are trained to exchange bidirectional information at multiple scales to improve learning. Third, up to this point, our framework can only handle two views. To extend it to multiple views, we intend to design a cross-multi-view LRSDL framework leveraging at the same time the richness and complementary of multiple MRI views. Specifically, we intend to combine different similarity patches obtained from different target views. Each view can also represent an MRI modality such as diffusion MRI (dMRI) or functional MRI (fMRI). Specifically, dMRI [43] modality has proved its effectiveness for the classification of cervical cancer, which could also have a significant impact on improving the tumor grade identification accuracy.

V. CONCLUSION

Cervical cancer is a recurrent gynecological malignancy and a communal cause of death. Cervical cancer staging is an invaluable tool in assessing the spread of the tumor to local and distant organs from the axial and sagittal views. In this paper, we proposed the first automated classification framework for cervical cancer grading. Specifically, we designed a CVSS dictionary learning framework (CVSS-LRSDL) using multi-view MR images to benefit from the unidirectional interaction from a source view to a target one for staging the disease. The focus of our paper is on the use of the SS patches for DL training. Our method achieved a significant increase in cancer grade classification accuracy when using both the advantages of the shared dictionary and the particular dictionary of the LRSDL approach. The experimental results, based on our MRI pelvic dataset, demonstrated the superior classification performance of our method, especially in determining the accurate stage. In our future work, we will first generalize our framework across multiple target MRI views or modalities such as dMRI or fMRI. Second, to improve learning, we will specify a bidirectional flow across views trading bidirectional information at multiple scales.

REFERENCES

[1] S. Pecorelli, "Revised figo staging for carcinoma of the vulva, cervix, and endometrium," *International Journal of Gynecology & Obstetrics*, vol. 105, no. 2, pp. 103–104, 2009.

[2] V. Nicolet, L. Carignan, F. Bourdon, and O. Prosmanne, "Mr imaging of cervical carcinoma: a practical staging approach," *Radiographics*, vol. 20, no. 6, pp. 1539–1549, 2000.

[3] Y. Okamoto, Y. O. Tanaka, M. Nishida, H. Tsunoda, H. Yoshikawa, and Y. Itai, "Mr imaging of the uterine cervix: imaging-pathologic correlation," *Radiographics*, vol. 23, no. 2, pp. 425–445, 2003.

[4] S. Choi and I.-C. J. Hsu, "Cervical cancer," in *Handbook of Evidence-Based Radiation Oncology*. Springer, 2018, pp. 623–644.

[5] N. Bnouni, O. Mechi, I. Rekik, M. S. Rhim, and N. E. B. Amara, "Semi-automatic lymph node segmentation and classification using cervical cancer mr imaging," in *Advanced Technologies for Signal and Image Processing (ATSIP), 2018 4th International Conference on*. IEEE, 2018, pp. 1–6.

[6] N. Bnouni, I. Rekik, M. S. Rhim, and N. E. B. Amara, "Dynamic multi-scale cnn forest learning for automatic cervical cancer segmentation," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2018, pp. 19–27.

[7] N. Bnouni, H. Ben Amor, I. Rekik, M. S. Rhim, B. Solaiman, and N. E. B. Amara, "Boosting cnn learning by ensemble image pre-processing methods for cervical cancer mr image segmentation." in *International conference on Sensors, Systems, Signals and advanced technologies (SSS), 2018*. IEEE, 2018.

[8] M. Aharon, M. Elad, A. Bruckstein *et al.*, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on signal processing*, vol. 54, no. 11, p. 4311, 2006.

[9] H. Zhang, N. M. Nasrabadi, Y. Zhang, and T. S. Huang, "Multi-view automatic target recognition using joint sparse representation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2481–2497, 2012.

[10] H. S. Mousavi, V. Monga, and T. D. Tran, "Iterative convex refinement for sparse recovery," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1903–1907, 2015.

[11] M. W. Spratling, "Image segmentation using a sparse coding model of cortical area v1," *IEEE transactions on image processing*, vol. 22, no. 4, pp. 1631–1643, 2013.

[12] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 895–900.

[13] X. Mo, V. Monga, R. Bala, and Z. Fan, "Adaptive sparse representations for video anomaly detection," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 4, pp. 631–645, 2014.

[14] T. Thongkamwitoon, H. Muammar, and P.-L. Dragotti, "An image recapture detection algorithm based on learning dictionaries of edge profiles," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 953–968, 2015.

[15] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification via kernel sparse representation," *IEEE Transactions on Geoscience and Remote sensing*, vol. 51, no. 1, pp. 217–231, 2013.

[16] X. Sun, Q. Qu, N. M. Nasrabadi, and T. D. Tran, "Structured priors for sparse-representation-based hyperspectral image classification," *IEEE geoscience and remote sensing letters*, vol. 11, no. 7, pp. 1235–1239, 2014.

[17] X. Sun, N. M. Nasrabadi, and T. D. Tran, "Task-driven dictionary learning for hyperspectral image classification with structured sparsity constraints," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 8, pp. 4457–4471, 2015.

[18] U. Srinivas, H. S. Mousavi, V. Monga, A. Hattel, and B. Jayarao, "Simultaneous sparsity model for histopathological image representation and classification," *IEEE transactions on medical imaging*, vol. 33, no. 5, pp. 1163–1179, 2014.

[19] T. H. Vu, H. S. Mousavi, V. Monga, G. Rao, and U. A. Rao, "Histopathological image classification using discriminative feature-oriented dictionary learning," *IEEE transactions on medical imaging*, vol. 35, no. 3, pp. 738–751, 2016.

[20] T. H. Vu, H. S. Mousavi, V. Monga, U. A. Rao, and G. Rao, "Dfdl: Discriminative feature-oriented dictionary learning for histopathological image classification," in *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*. IEEE, 2015, pp. 990–994.

[21] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

[22] L. Li, S. Li, and Y. Fu, "Learning low-rank and discriminative dictionary for image classification," *Image and Vision Computing*, vol. 32, no. 10, pp. 814–823, 2014.

[23] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Sparse representation based fisher discrimination dictionary learning for image classification," *International Journal of Computer Vision*, vol. 109, no. 3, pp. 209–232, 2014.

[24] Z. Jiang, Z. Lin, and L. S. Davis, "Label consistent k-svd: Learning a discriminative dictionary for recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 11, pp. 2651–2664, 2013.

[25] ——, "Learning a discriminative dictionary for sparse coding via label consistent k-svd," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1697–1704.

[26] Q. Zhang and B. Li, "Discriminative k-svd for dictionary learning in face recognition," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2691–2698.

[27] E. L. Bases, "Discriminative feature-oriented dictionary learning for histopathological image classification."

[28] S. Gao, I. W.-H. Tsang, and Y. Ma, "Learning category-specific dictionary and shared dictionary for fine-grained image categorization," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 623–634, 2014.

[29] I. Ramirez, P. Sprechmann, and G. Sapiro, "Classification and clustering via dictionary learning with structured incoherence and shared features," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 3501–3508.

[30] S. Kong and D. Wang, "A dictionary learning approach for classification: separating the particularity and the commonality," in *European Conference on Computer Vision*. Springer, 2012, pp. 186–199.

[31] N. Zhou and J. Fan, "Jointly learning visually correlated dictionaries for large-scale visual recognition applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 4, pp. 715–730, 2014.

[32] T. H. Vu and V. Monga, "Fast low-rank shared dictionary learning for image classification," *IEEE Transactions on Image Processing*, vol. 26, no. 11, pp. 5160–5175, 2017.

[33] ——, "Learning a low-rank shared dictionary for object classification," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 4428–4432.

[34] M. Yang, L. Zhang, X. Feng, and D. Zhang, "Fisher discrimination dictionary learning for sparse representation," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 543–550.

[35] P. L. Suárez, A. D. Sappa, and B. X. Vintimilla, "Cross-spectral image patch similarity using convolutional neural network," in *Electronics, Control, Measurement, Signals and their Application to Mechatronics (ECMSM), 2017 IEEE International Workshop of*. IEEE, 2017, pp. 1–5.

[36] T. Deselaers and V. Ferrari, "Global and efficient self-similarity for object classification and detection," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1633–1640.

[37] J. Xu, L. Zhang, W. Zuo, D. Zhang, and X. Feng, "Patch group based nonlocal self-similarity prior learning for image denoising," in *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 244–252.

[38] C. Cruz, R. Mehta, V. Katkovnik, and K. O. Egiazarian, "Single image super-resolution based on wiener filter in similarity domain," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1376–1389, 2018.

[39] A. Bustin, D. Voilliot, A. Menini, J. Felblinger, C. de Chillou, D. Burschka, L. Bonnemains, and F. Odille, "Isotropic reconstruction of mr images using 3d patch-based self-similarity learning," *IEEE Transactions on Medical Imaging*, 2018.

[40] J. V. Manjón, P. Coupé, A. Buades, D. L. Collins, and M. Robles, "Mri superresolution using self-similarity and image priors," *Journal of Biomedical Imaging*, vol. 2010, p. 17, 2010.

[41] F. Pozzi, T. Di Matteo, and T. Aste, "Exponential smoothing weighted correlations," *The European Physical Journal B*, vol. 85, no. 6, p. 175, 2012.

[42] S. Amiri, M. A. Mahjoub, and I. Rekik, "Dynamic multiscale tree learning using ensemble strong classifiers for multi-label segmentation of medical images with lesions," *Neurocomputing*, 2018.

[43] A. F. A. Ghany, M. A. A. Wahed, and W. K. AL-Habash, "Role of magnetic resonance imaging in diagnosis and staging of uterine cervical carcinoma." *Egyptian Journal of Hospital Medicine*, vol. 71, no. 5, 2018.