A MULTIPLE COVARIANCE APPROACH FOR CELL DETECTION OF GRAM-STAINED SMEARS IMAGES

Matthew Crossman[⋆]*, Arnold Wiliem*[⋆]*, Paul Finucane*[†]*, Anthony Jennings*[†]*, and Brian C. Lovell*[⋆]

[⋆]The University of Queensland, Australia
[†]Sullivan Nicolaides Pathology, Australia

ABSTRACT

Microscope examination of Gram stained clinical specimens is used for aiding the diagnosis of patients with infectious diseases. In high volume pathology laboratories, this manual microscopy examination is considered time consuming and labour intensive. Unfortunately, despite the great benefits offered from the application of Computer Aided Diagnosis (CAD) systems, to our knowledge, the highest automation stage for Gram stained slide analysis is only at the pre-analytical process. This paper takes the first steps towards the application of computer vision to direct smear, Gram stained images. To that end, we present a novel Gram stain image dataset. In addition, we also propose a multiple covariance approach for leukocyte and epithelial cell detection in Gram stain images. Each covariance matrix represents a particular image region characterising the cell's deformed structure. As covariance matrices form points on an Symmetric Positive Definite (SPD) manifold, the traditional Euclidean-based analysis cannot be used. As such, we first map the manifold points into the Reproducing Kernel Hilbert Space (RKHS). The analysis is done via a novel kernel similarity function that allows comparison between sets of covariance matrices. The proposed approach is contrasted, in the proposed dataset, with two recent state of the art methods in pedestrian detection: Histogram Of Gradient (HOG) and the traditional single covariance matrix approach. We found that the proposed approach outperformed both of these methods.

Index Terms— Gram stain analysis, direct smears, Cell detection, Riemannian manifolds, Symmetric Positive Definite Matrix group

1. INTRODUCTION

Recently there has been growing interest in applying image analysis to pathology test images [1, 2, 3, 4]. More precisely, Computer Aided Diagnosis (CAD) systems were developed to automatically provide analysis based on the input images. Results produced by these methods can be used to support the scientists' manual/subjective analysis; making the test results more reliable and consistent. In microbiology, it is known that microscopic examination of Gram stained preparations of clinical specimens is valuable for physicians who are managing patients with infectious diseases [5]. This is due to the fact that it is a rapid and cost effective procedure.

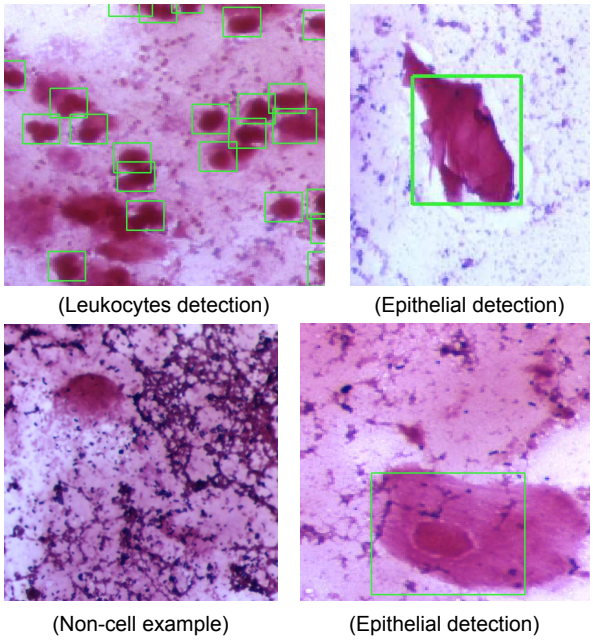

Fig. 1. Some detection bounding box results from our proposed approach. Note that the cells of interest such as leukocytes and epithelial cells have extreme variability in shapes. In some cases, some non-cell objects also have similar appearance to these cells. Our proposed approach is able to address these variabilities.

The Gram staining protocol is a staining procedure applied to tissue samples on glass slides, which facilitates classification of cells under bright field microscopes as either Gram positive or Gram negative [6]. This classification aids the scientist in estimating the abundance of several cell classes in the smear. The relative population of white blood cells and bacteria cells, together with epithelial cells gives an efficient first diagnosis of bacterial infections.

For pathology laboratories handling large volumes of patient samples, the manual examination of each slide significantly consumes valuable time of scientists. These issues could be addressed using CAD systems designed for cell recognition with the detection statistics used for population

This research was funded by Sullivan Nicolaides Pathology (SNP), Australia and the Australian Research Council (ARC) Linkage Projects Grant LP130100230.

estimation. Unfortunately, despite the significant benefits offered by CAD systems, to the best of our knowledge the highest automation stage for Gram stained slide analysis is only at the pre-analytical process [6].

To that end, in this study we present an application of computer vision to the area of Gram stained direct smear images. In particular, we confine ourselves to the problem of epithelial and white blood cell detection in Gram stain images. Whilst, the white blood cells (leukocytes) are an important parameter to infer possible bacterial infection, the epithelial cells are considered for determining the specimen quality. A large number of white blood cells could indicate that the patient is having a severe infection. A large concentration of white blood cells could also be used as a proxy to search for bacteria which has a significantly smaller size. On the other hand, a large number of epithelial cells could mean that the specimen may not have been properly taken.

In this work, we propose a novel benchmarking Gram stain direct smear image dataset which has three benefits: (1) To generate interest within the community and thus advance the field; (2) To assist the development of novel computer vision techniques to address unique problems presented in this application domain; (3) To provide a general and meaningful platform for practitioners in this field for comparing their algorithms.

In the second contribution, we present evaluation results on two state-of-the-art of pedestrian detection methods. We opt to use pedestrian detection as cells, like pedestrians, can be considered as deformable objects. One of the differences between cells and pedestrians is that for cells, the within class deformation has significant variability with respect to relative positioning and orientation. Fig. 1 presents some sample images and results of our proposed approach. This is especially true for Gram stained direct smear images due to the fact that when the smears are prepared the scientists manually swab the specimen on the glass slide.

In light of this fact, we argue that it is more appropriate to address this problem by modelling the cell images as covariance matrices, which can be interpreted as points over the Symmetric Positive Definite (SPD) manifolds. The Covariance of features approach has been proven to be effective in pedestrian detection [7]. As shown in [7], this modelling is more powerful than the Euclidean features such as Histogram Oriented Gradient (HOG) features. Unfortunately, we found that the existing approaches, mainly focussing on a single covariance features (*i.e.* one image is represented with one covariance matrix), is not adequate for our problem due to the high within class variability in the cells of interest. As such we propose a multiple covariance approach which encodes much richer information to characterise a cell. In our evaluation we found that the proposed multiple covariance approach outperformed both the traditional single covariance approach [7] and one of the state-of-the-art approaches for pedestrian detection [8].

We continue our paper as follows. We first described the task in Section 2. The proposed benchmarking platform is described in Section 3. Section 4 elaborates on our proposed multiple covariance approach. We present our evaluation results in Section 5. Section 6 discusses main findings and future directions.

2. PROBLEM DESCRIPTION

Given an image $\boldsymbol{I}$, the goal is to classify the image into the following classes $\mathcal{C} = \{\text{non-cell}, \text{epithelial}, \text{leuokocyte}\}$. Note that the number of non-cell exemplars is much higher than the other two classes. This means, we could pose the problem in two different forms: (1) Cell vs non-cell detection; (2) Epithelial cell detection and Leukocyte detection. In this work, we consider these two forms.

3. GRAM-STAINED SMEARS IMAGE DATASET

To address the evaluation of different recognition methods on Gram stained smears, the SNP-Gramstain dataset was created. The dataset contains 150 colour images of anonymous Gram stained smears that were collected at Sullivan Nicolaides Pathology (SNP) Australia. The dataset will be available at `http://www.itee.uq.edu.au/sas/datasets`.

The images were taken using a Pixel link coloured camera mounted on a motorised microscope at x63 magnification and with a resolution of 2048x1536 pixels in uncompressed TIF format. From the 150 images, positive and negative examples of leukocyte and epithelial cells were cropped, with some examples shown in Fig. 1. The 8129 cropped examples are composed of: (1) 586 leukocytes; (2) 162 epithelial cells; (3) 2471 hand-labelled negative exemplars and (4) 4910 automatically extracted negative exemplars. The automatically generated negative exemplars were extracted from 11 of the dataset images, which contained no cells of interest. From these images, square crops were extracted in a grid pattern at a wide range of scales, with wider spacing between larger windows to reduce overlap. The hand labelled negative examples were selected to present difficult training and testing exemplars. Negative examples focus on artefacts or background materials that classifiers might have difficulty with. The negative examples also contain leukocyte sized windows taken on the border of large epithelial cells, to help train and test classifiers ability to avoid false-positive leukocyte detections.

4. PROPOSED APPROACH

As mentioned, our proposal is based on the work proposed in [7]. The work represents an image by a positive definite covariance matrix. Covariance matrices belong to the Symmetric Positive Definite (SPD) group. These matrices can be interpreted as points over Riemannian manifolds; thus we name the manifolds as SPD manifold. We first present an overview of the SPD manifold, the traditional single covariance matrix approach, and then our proposed approach is described.

4.1. Symmetric Positive Definite Manifold

Riemannian manifolds are smooth and differentiable manifolds embedded in a higher dimensional space. In general, the manifold topology is assumed to be either unknown or complex. Thanks to the fact that Riemannian manifolds are smooth and differentiable, it becomes possible to study their

surface by using differential geometry tools. In other words, one can derive the function mapping from a location on the manifold onto its corresponding tangent space by studying the derivative operator on the location. In fact, the geodesic distance (*i.e.* the true distance between two points on the manifold) is generally derived via tangent space. We refer interested readers to [9, 7] for a full treatment of Riemannian manifolds.

The geodesic distance between two points $\boldsymbol{X}_1, \boldsymbol{X}_2 \in \mathrm{SYM}_d$ in the SPD manifold is defined via [7].

$$d_g(\boldsymbol{X}_1, \boldsymbol{X}_2) = \mathrm{trace}\left\{\mathrm{logm}^2(\boldsymbol{X}_1^{-\frac{1}{2}}\boldsymbol{X}_2\boldsymbol{X}_1^{-\frac{1}{2}})\right\}^{\frac{1}{2}} \quad (1)$$

where $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are the $d \times d$ SPD matrices; $\mathrm{logm}(\cdot)$ is the matrix logarithm.

Often it is useful to study the SPD manifold by first mapping the points into the Reproducing Kernel Hilbert Space (RKHS) to significantly improve the accuracy [10, 11, 12]. In this work, we opt to use the Log-Euclidean kernel due to its high accuracy and low computational complexity [12].

$$k_e(\boldsymbol{X}_1, \boldsymbol{X}_2) = \| \mathrm{logm}(\boldsymbol{X}_1) - \mathrm{logm}(\boldsymbol{X}_2)\|_{\mathrm{F}} \quad (2)$$

where $k_e(\cdot)$ is the Log-Euclidean kernel similarity function; and $\| \cdot \|_{\mathrm{F}}$ is the Frobenious norm. Once the manifold points have been mapped into the RKHS, we employ the Kernel SVM to build the detector.

4.2. Single Covariance Approach

To represent an image $\boldsymbol{I}$ as a covariance matrix $\boldsymbol{X}$, for each pixel location in $\boldsymbol{I}$ we first extract an 11 dimensional feature vector composed by.

$$\left[u, v, \boldsymbol{I}^r_{u,v}, \boldsymbol{I}^g_{u,v}, \boldsymbol{I}^b_{u,v}, \left|\frac{\partial \boldsymbol{I}^r_{u,v}}{\partial u}\right|, \left|\frac{\partial \boldsymbol{I}^r_{u,v}}{\partial v}\right|, \cdots, \left|\frac{\partial \boldsymbol{I}^b_{u,v}}{\partial u}\right|, \left|\frac{\partial \boldsymbol{I}^b_{u,v}}{\partial v}\right|\right] \quad (3)$$

where u, v are the pixel coordinates; $\boldsymbol{I}^r_{u,v}$ is the grey value of red channel at location (u, v); $\boldsymbol{I}^g_{u,v}$ is the grey value of green channel at location (u, v); $\boldsymbol{I}^b_{u,v}$ is the grey value of blue channel at location (u, v); $\left|\frac{\partial \boldsymbol{I}^r_{u,v}}{\partial v}\right|$ is value of the first image derivative w.r.t v axis at location (u, v).

The 11×11 covariance matrix $\boldsymbol{X}$ is then computed once the feature vectors are extracted from every pixel location. This becomes the representation of $\boldsymbol{I}$ over the SPD manifold space.

4.3. Multiple Covariance Approach

Instead of extracting only one covariance matrix, we first divide an image into three regions characterising a cell: (1) inner region which encodes the cell content; (2) outer region that contains the information between the cell cytoplasm and background; and (3) the whole region which includes region (1) and (2). More precisely, given an image $\boldsymbol{I}$, we set the inner region to be τ times of the size $\boldsymbol{I}$, $0 < \tau < 1$. For instance if the size of $\boldsymbol{I}$ is 100×64 and $\tau = 0.5$, then the inner region size is 50×32. The outer region is the remaining region in the image that does not belong to the inner region.

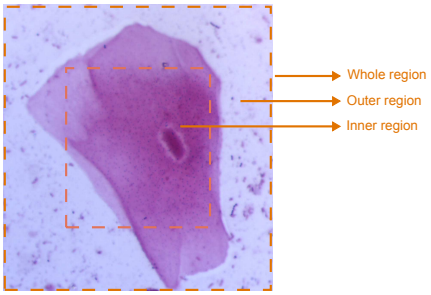

Fig. 2. The illustration on how the regions are divided. Inner region is always located at the center with size τ times smaller than the whole region.

Fig. 2 sketches how the three regions are divided. Our proposed approach is inspired from the Cell Pyramid Matching (CPM) descriptor which was proposed to address cell classification in fluorescence images [1]. Unlike the CPM descriptor however, we extract a covariance descriptor from each region. Let $\{\boldsymbol{X}^r_1\}_{r=1}^{r=3}, \{\boldsymbol{X}^r_2\}_{r=1}^{r=3}$ be the three covariance features extracted from the three regions of images $\boldsymbol{I}_1$ and $\boldsymbol{I}_2$, respectively. We note that it is not trivial to determine the similarity between these two points. Henceforth, we propose the modification of Log-Euclidean kernel as follows.

$$k_{me}(\{\boldsymbol{X}^r_1\}_{r=1}^{r=3}, \{\boldsymbol{X}^r_2\}_{r=1}^{r=3}) = \|\phi(\{\boldsymbol{X}^r_1\}_{r=1}^{r=3}) - \phi(\{\boldsymbol{X}^r_2\}_{r=1}^{r=3})\|_{\mathrm{F}} \quad (4)$$

where $k_{me}(\cdot)$ is the proposed modified Log-Euclidean and $\phi(\cdot)$ is defined as the function that concatenates the matrix logarithms of individual region.

$$\phi(\{\boldsymbol{X}^r_i\}_{r=1}^{r=3}) = \left[\mathrm{logm}(\boldsymbol{X}^1_i)\, \mathrm{logm}(\boldsymbol{X}^2_i)\, \mathrm{logm}(\boldsymbol{X}^3_i)\right] \quad (5)$$

As derived from the original Log-Euclidean kernel, the kernel k_{me} is always symmetric and positive semi-definite, thus it is a Mercer Kernel.

5. EVALUATION OF RESULTS

We contrasted our proposed approach, denoted Multiple Covariance with Kernel SVM (MC-KSVM), to two recent state-of-the-art approaches in the pedestrian detection domain: (1) the Histogram of Oriented Gradient(HOG) with Linear SVM [8], denoted HOG-SVM and (2) the Single Covariance approach in conjunction with Kernel SVM (SC-KSVM) [7]. We note, that for a fair comparison, we used Kernel SVM for SC-KSVM with kernel similarity function defined in Eqn. 2. Both SC-KSVM and MC-KSVM employed the same features described in Eqn. 3. Therefore, the only difference between SC-KSVM and MC-KSVM is the fact that MC-KSVM employs multiple covariance matrices to represent an image.

The detection rates reported on the dataset were all recorded with a False Positive Rate (FPR) of 10^{-3} across the dataset. Note that, here cropped cell images are used. Thus, the performance is not computed based on the overlap of the bounding boxes. As the dataset contains a significant fraction of difficult negative exemplars, the false positive rate is expected to remain relevant when performing sliding window detection in the full non-cropped images. Evaluation was done using 10-fold cross validation on each of the labelled cropped images. We followed the approach used in [8]

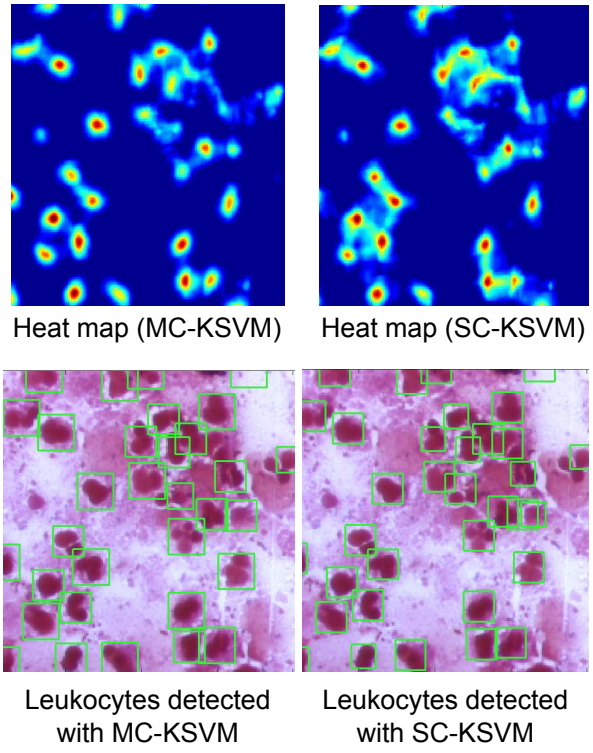

Fig. 3. Comparisons between MC-KSVM (proposed) and SC-KSVM on leukocyte detection.

Method	Epithelial	Leukocytes	Both
HOG-SVM [8]	61.3%	65.5%	50.7%
SC-KSVM [7]	71.3%	87.9	79%
MC-KSVM(Proposed)	**89.6%**	**90.4%**	**86.6%**

Table 1. Detection percentage at 10^{-3} False positive Rate. HOG-SVM: Histogram Oriented Gradient with SVM; SC-KSVM: Single Covariance approach with Kernel SVM, please refer to text for detail descriptions; MC-KSVM: Multiple Covariance approach with KSVM (proposed approach).

in order to produce the detection rates with the designated FPR. For the sake of brevity, from now on, we will refer to "detection rates with FPR of 10^{-3}" as simply the "detection rates". All the hyperparameters were determined from the cross-validation set.

Table 1 shows the detection rate of each detection method, for the detection of leukocytes, epithelial and both types of cell together. It can be seen that the multiple covariance method achieves the highest detection rates for every cell type. HOG-SVM achieves comparatively poor performance on the recognition task. The performance of HOG-SVM can be attributed to the highly variable shape of the cells, where HOG attempts to learn a rotationally variant structure. Single covariance SC-SVM features achieve respectable performance. However its performance is limited to the noisy environments where the features cannot adequately encode the cells.

Fig. 3 shows an example of the sliding window SVM

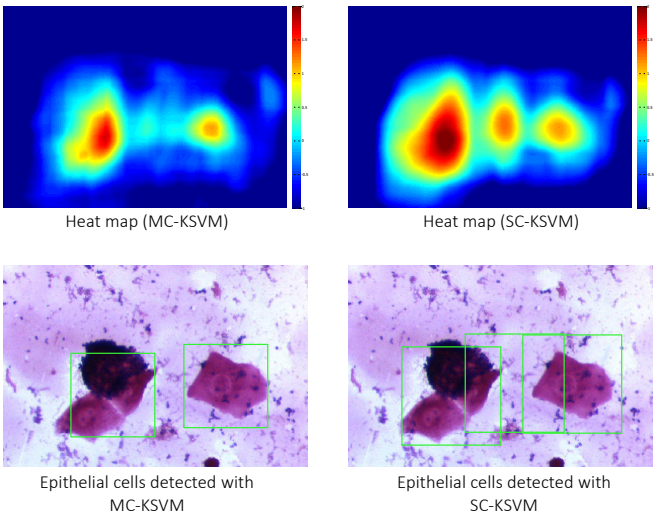

Fig. 4. Comparisons between MC-KSVM (proposed) and SC-KSVM on epithelial cell detection.

confidence heatmap and detection bounding box for leukocyte detection using SC-SVM and MC-KSVM methods. By visual examination, the MC-KSVM method yields a cleaner heatmap than SC-KSVM, with inter cell detections suppressed. Fig. 4 shows an Epithelial detection example, similar to Fig. 3. For epithelial detection, the improvement of MC-KSVM over SC-KSVM is more significant. The SC-KSVM heatmap has a strong spurious detection in the space between epithelial cells, that is not present in the MC-KSVM heatmap. This increased difference is attributed to the greater variability in epithelial cell shape, where the additional structure of MC-KSVM allows for clearer discrimination.

6. MAIN FINDINGS AND FUTURE DIRECTION

This work tackles the problem of Leukocyte and Epithelial cell detection on Gram-stained smear images. To the best of our knowledge, this is one of the first attempts to push the automation stage in this field, which is currently only at the pre-analytical stage. We first create a benchmarking dataset for this purpose. Then, a novel cell detection approach was proposed, making use of multiple covariance matrices for representing a cell image. As covariance matrices form a manifold space in the SPD manifolds, one cannot use the traditional Euclidean-based techniques. Therefore, we opt to use kernel analysis. To that end, a novel kernel similarity function which measures the similarity between two set of multiple covariance matrices is proposed. We contrasted the proposed approach to recent state-of-the-art approaches in pedestrian detection: HOG-SVM and the traditional single covariance approach. Evaluation results found the proposed method performed best on each of the cell classes tested. Future work aims to expand the dataset to include a greater quantity of samples and to extend the detection analysis to the bacterial cells. We also plan to further reduce the computational complexity using random projection methods described in [13, 14]

7. REFERENCES

[1] Arnold Wiliem, Conrad Sanderson, Yongkang Wong, Peter Hobson, Rodney F. Minchin, and Brian C. Lovell, "Automatic classification of human epithelial type 2 cell indirect immunofluorescence images using cell pyramid matching," *Pattern Recognition*, vol. 47, no. 7, pp. 2315 – 2324, 2014.

[2] Arnold Wiliem, Peter Hobson, Rodney F. Minching, and Brian C. Lovell, "A bag of cells approach for antinuclear antibodies hep-2 image classification," *Cytometry Part A. In press.*, 2015.

[3] Yan Yang, Arnold Wiliem, Azadeh Alavi, Brian C. Lovell, and Peter Hobson, "Visual learning and classification of human epithelial type 2 cell images through spontaneous activity patterns," *Pattern Recognition*, vol. 47, no. 7, pp. 2325–2337, 2014.

[4] Paul J Tadrous, "Computer-assisted screening of ziehl-neelsen-stained tissue for mycobacteria. algorithm design and preliminary studies on 2,000 images," *American Journal of Clinical Pathology*, vol. 133, no. 6, pp. 849–858, June 2010.

[5] Linda M. Marler, Jean A. Siders, and Stephen Allen, *Direct Smear Atlas: A Monograph of Gram-Stained Preparations of Clinical Specimens*, Lippincott Williams and Wilkins, 2001.

[6] E. Baron, S. Mix, and W. Moradi, "Clinical utility of an automated instrument for gram staining single slides," *Journal of Clinical Microbiology*, vol. 48, no. 6, pp. 2014–2015, 2010.

[7] Oncel Tuzel, Fatih Porikli, and Peter Meer, "Pedestrian detection via classification on Riemannian manifolds," *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, vol. 30, no. 10, pp. 1713–1727, 2008.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[9] Xavier Pennec, "Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements," *Journal of Mathematical Imaging and Vision*, vol. 25, no. 1, pp. 127–154, 2006.

[10] Mehrtash Harandi, Conrad Sanderson, Arnold Wiliem, and Brian C. Lovell, "Kernel analysis over riemannian manifolds for visual recognition of actions, pedestrians and textures," in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2012.

[11] Mehrtash Harandi, Conrad Sanderson, Sareh Shirazi, and Brian C. Lovell, "Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[12] S. Jayasumana, Richard Hartley, M. Salzmann, H. Li, and M. Harandi, "Kernel methods on riemannian manifold of symmetric positive definite matrices," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[13] Azadeh Alavi, Arnold Wiliem, Kun Zhao, Brian C Lovell, and Conrad Sanderson, "Random projections on manifolds of symmetric positive definite matrices for image classification," in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2014, pp. 301–308.

[14] Kun Zhao, Arnold Wiliem, and Brian C. Lovell, "Kernelised orthonormal random projection on grassmann manifolds with applications to action and gait-based gender recognition," in *IEEE Conference on Identity, Security and Behaviour Analysis*, 2015.