Tensor Decomposition of Largest Convolutional Eigenvalues Reveals Pathological Predictive Power of RhoB in Rectal Cancer Biopsy

Tuan Pham (✉ tdpham123@gmail.com)

Prince Mohammad Bin Fahd University

Vinayakumar Ravi

Prince Mohammad Bin Fahd University

Na Liu

Linköping University

Bin Luo

Linköping University

Chuan-Wen Fan

Linköping University

Xiao-Feng Sun

Linköping University

Research Article

Keywords: Artificial Intelligence, Rho GTPase family, radiotherapy (RT), RhoA, RhoB, RhoD, Gastrointestinal Surgery

Posted Date: October 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-957359/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Tensor Decomposition of Largest Convolutional Eigenvalues Reveals Pathological Predictive Power of RhoB in Rectal Cancer Biopsy

Tuan D. Pham^{1,*}, Vinayakumar Ravi¹, Na Liu², Bin Luo^{2,3}, Chuan-Wen Fan², and Xiao-Feng Sun^{2,*}

¹Prince Mohammad Bin Fahd University, Center for Artificial Intelligence, Khobar 31952, Saudi Arabia

²Linköping University, Department of Oncology and Department of Biomedical and Clinical Sciences, 58185 Linköping, Sweden

³Sichuan Provincial People's Hospital, Department of Gastrointestinal Surgery, Chengdu, China

*Corresponding authors (emails: tpham@pmu.edu.sa; xiao-feng.sun@liu.se)

ABSTRACT

RhoB protein belongs to the Rho GTPase family, which plays an important role in governing cell signaling and tissue morphology. RhoB expression is known to have implications in pathological processes of diseases. Investigation in the regulation and communication of this protein detected by immunohistochemical staining on the microscope is worth exploring to gain insightful information that may lead to identifying optimal disease treatment options. In particular, the role of RhoB in rectal cancer is not well-discovered. Here, we report that methods of deep learning-based image analysis and the decomposition of multiway arrays discover the predictive factor of RhoB in two cohorts of rectal-cancer patients having survival rates of less and more than 5 years. The analysis results show distinctions between the tensor-decomposition factors of the two cohorts.

Introduction

Rectal cancer is one of the most common cancers worldwide¹. Preoperative radiotherapy (RT) is a standard neoadjuvant or palliative therapy for patients with advanced cancer, but many cancers have no or little response to RT, leading to patients' poor prognosis²⁻⁴. One of the main reasons is that currently used clinicopathological factors, even tumor stage, cannot precisely provide evidence for clinicians to design an efficient RT strategy. Thus, it is important to find promising biological factors that can be used to predict a cancer patient's response to a specific treatment, or the patient's outcome of developing the disease.

The Rho protein family plays a key role in signal transduction or cell signaling that regulates the actin cytoskeleton and phagocyte NADPH (nicotinamide adenine dinucleotide phosphate) oxidase⁵. It was found that RhoA, RhoB, and RhoD have effects on the membrane trafficking. However, knowledge about the underlying molecular mechanisms for these influences of Rho proteins is incomplete⁶. Growing interests in studying the role of RhoB in cancer have been reported in literature⁷⁻¹³.

RhoB is a distinct anti-apoptotic protein. It was demonstrated that the p53 isoform $\delta 133p53\beta$ involves in the regulation of the apoptotic response in colorectal cancer cell lines¹⁴. The role of RhoB protein in radioresistance of colorectal cancer has recently been investigated¹⁵. This work found that the overexpression of RhoB is associated with poor overall survival in rectal-cancer patients undergoing radiotherapy (RT), indicating the prognostic factor of the protein in relation to RT. However, the work failed to establish evidence of the predictive factor of RhoB in rectal-cancer patients before RT.

By using the basis of feature transformation developed for deep learning in artificial intelligence (AI) and multidimensional data analysis, the study addressed in this paper presents for the first time that RhoB expression on immunohistochemistry (IHC) imaging obtained from biopsy samples of rectal-cancer patients without RT can differentiate groups between less and more than 5 years of overall survival. Without relying on subjective assessment of the protein expression, the current approach objectively extracts hidden and deep features of the biopsy samples from different three color-image channels and integrates the information from multiple samples and multiple channels by means of a method for tensor decomposition that is viewed as extended principal component analysis. Tensor decomposition is capable of discovering unique latent components from complex data of multiple dimensions¹⁶.

In fact, deep learning in AI has been recently realized as a very useful method for characterizing nanobiological structures on microscopy and much more effective approach than traditional image analysis algorithms¹⁷. The work reported in¹⁸ used deep learning and Gaussian mixture modeling to characterize variability in cryogenic electron microscopy and determine

protein conformational landscapes. The application of deep neural networks presented in the work had certain advantages over conventionally adopted methods, such as the manifold learning. A study on imaging cellular nanostructures addressed in¹⁹ introduced a deep-learning model to overcome a limitation encountered with standard methods. The results showed that the utilization of deep learning could accurately localize single emitters at high density in three dimensions for a variety of imaging modalities and conditions.

Tensor decomposition^{20,21} has been increasingly applied for solving challenging problems in medicine and health over recent years^{22–28}. In cancer, tensor decomposition has recently been applied for identifying microRNA (miRNA) and mRNA expression profiles as potential prognostic biomarkers for kidney renal clear cell carcinoma, where genes involving in cancer-related pathways were found to be associated with miRNAs²⁹. Most recently, tensor decomposition has been reported as a useful computational data analysis tool that provides an effective way for integrative analysis of epigenomic data. The tensor decomposition and classification of epigenomic tensors could differentiate epigenomic features between normal myometrium and leiomyoma subtypes, and identified the HOXA13 gene as a potential tumorigenic factor in leiomyoma³⁰.

Based on promising biomedical results recently achieved with applications of advanced AI and data science tools, it is important to explore these methods for discovering disease biomarkers, including molecular, histologic, physiologic, and radiative attributes. Here, we show that the novel combination of deep learning-based feature extraction of RhoB expression on IHC imaging obtained from the Swedish Rectal Cancer Trial³¹, and the tensor decomposition of these features in a multidimensional setting can reveal the predictive factor of this protein in rectal-cancer patients.

Methods

Patient samples

This study included 91 patients (51 male and 40 female) who were of the South-East Swedish Health Care region and participated in the randomized Swedish Rectal Cancer Trial of preoperative RT between 1987 and 1990³¹. Each participant signed the informed consent. Among the patients, 38 were subjected to 25Gy radiotherapy in 5 fractions and followed by surgical operation within 2 weeks (median 6 days, 0-13 days), and the rest of 53 patients had operation alone. None of the patients received preoperative or adjuvant chemotherapy. Two groups had no statistical differences regarding clinical and pathological characteristics such as the gender, age, TNM stage, grade of differentiation etc ($p < .05$). All patients were followed up to more than 5 years. All the 120 biopsy samples (1-3 samples from each of 91 patients) were taken before operation, in which 45 and 75 tissue microarray biopsy samples of IHC staining obtained from rectal-cancer patients having less and more than 5-year disease-free survival, respectively.

The study was conducted in accordance with the Declaration of Helsinki. The experimental protocols were approved by the Institutional Ethics Committee of Linköping University, Sweden (Dnr-2012-107-31). The informed consent was signed by each participant.

Immunohistochemistry (IHC)

The biopsy samples were fixed with formalin, embedded in paraffin, and made as tissue microarray (TMA). TMA sections (4 μm) were deparaffinized by 100% xylene, incubated in 100% ethanol, and rehydrating with decreasing concentrations of ethanol. After antigen retrieval, the sections were incubated with anti-RhoB mouse monoclonal (sc-8048) overnight, and then incubated with a secondary antibody, Envision System Labelled Polymer-HRP Anti-Mouse for 25 minutes. After reacting with Liquid DAB+ (Dako), the sections were lightly counterstained with hematoxylin. The slides were scanned using Leica Aperio CS2 scanner.

Figure 1 shows some IHC biopsy samples of RhoB expression obtained from rectal cancer patients whose survival rates were less or more than 5 years. Image sizes are about $3000 \times 3000 \times 3$. To discover the predictive factor of RhoB, the image color is conventionally examined by pathologists to determine if the expression is either negative or positive, or alternatively either weak, medium, or strong. However, due to the complexity of texture and color distribution of these images, human expert perception can be biased and subject to a high degree of uncertainty in assessing the expression.

Convolutional eigenvalues of an image

Using the idea of convolutional neural networks in deep learning, a sequence of maximum convolutional eigenvalues of a 2-D image or a color-image channel is computed in this study and considered as a kind of “deep” feature of the image. Convolutional eigenvalues of a fuzzy recurrence plot (FRP) has recently been introduced in literature³². The main difference between the convolutional eigenvalues of an image and an FRP is that the former extracts sequential eigenvalues of the corresponding convolved images, whereas the later compute the eigenvalues of the final convolved matrix of an FRP. The motivation for extracting sequential convolutional eigenvalues of an image is to capture complex RhoB expression on IHC. A sequence of maximum eigenvalues can extract complex image features from low to high levels while the iterative process of applying a kernel filter on the image is carried out. These features include intensity distribution, edges, and texture in the early process

of image convolution, and more abstract properties of the image in later process; which can provide distinctive features for effective pattern discrimination.

An image convolution is performed by multiplying the values of a pixel and its neighbors with a small matrix, which is called a convolution matrix, mask, or kernel, for achieving different purposes, such as smoothing, sharpening, enhancing, and edge detection.

The convolution of a kernel $k(m,n)$ and an image $f(x,y)$ results in a filtered or convolved image $g(x,y)$, which can be generally expressed as

$$\begin{aligned} g(x,y) &= k(m,n) \circledast f(x,y) \\ &= \sum_{m=-a}^a \sum_{n=-b}^b k(m,n) f(x+m, y+n), \end{aligned} \quad (1)$$

where $\circledast$ denotes the convolution operator.

A desire is to emphasize transitions in the image intensity of the protein expression. Therefore, the Laplacian kernel that uses the second derivative for image sharpening was adopted in this study. The Laplacian operator, which takes partial derivatives along the two spatial axes of a 2-D image $f(x,y)$, is defined as³³

$$\begin{aligned} \nabla^2 f &= \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \\ &= f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y), \end{aligned} \quad (2)$$

which highlights sharp intensity transitions and reduces the effect of regions having slowly varying gray levels, resulting in the following 3×3 Laplacian mask

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3)$$

Furthermore, to rectify the effect of featureless background while still keeping the sharpened result, this can be done by subtracting the original from Laplacian image as follows.

$$\begin{aligned} f^*(x,y) &= f(x,y) - \nabla^2 f \\ &= f(x,y) - [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y)], \\ &= 5f(x,y) - [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)], \end{aligned} \quad (4)$$

which yields the 3×3 composite sharpening kernel

$$k = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (5)$$

By resizing the image to a square image of a certain size and after the convolution, the rectified linear unit (ReLU) is applied to eliminate negative values of the convolved image, and then the largest eigenvalue of the square image is computed. The ReLU function returns either a zero or positive value to each filtered pixel as follows.

$$g^*(x,y) = \max[0, g(x,y)]. \quad (6)$$

The next step is to reduce the size of the convolved image while still keeping useful information by using a pooling operator. There are different types of image pooling, such as the max, average, and sum operator. In deep learning, the most widely used pooling is the max pooling operator. Mathematically, let $\Omega = (\Omega_1, \Omega_2, \dots, \Omega_Q)$, where $\Omega_q = (\omega_{q,1}, \omega_{q,2}, \dots, \omega_{q,m \times m})$, $q = 1, 2, \dots, Q$, be a set of pooling regions. The number of pooling regions Q within a convolved image is determined by the pool size $m \times m$ and stride. The max pooling operating on a pooling region of size $m \times m$, denoted as $P_{\max}$, is defined as

$$P_{\max}(\Omega_q) = \max_{1 \leq l \leq m \times m} (\omega_{q,l}). \quad (7)$$

The procedure is then repeated with the process of convolution, ReLU, calculation of the largest eigenvalue, and pooling until the convolved image reaches to a specified size. The procedure is outlined with the steps as follows.

Procedure for Computing a Sequence of Largest Eigenvalues of an Image:

1. Input: An $N \times N$ image or resize the image to $N \times N$ size if it is not a square.
2. Given w , ReLU, pool size, and stride.
3. Given n as the stopping value for the iterative convolution.
4. While $N > n$
5. Perform the image convolution.
6. Apply ReLU on the convolved image.
7. Determine the largest eigenvalue of the convolved and ReLU-operated image.
8. Perform max pooling on the convolved and ReLU-operated image.
9. End While loop.

Figure 2 (A) shows the flowchart for computing a sequence of largest convolutional eigenvalues of a color-image channel.

PARAFAC (parallel factor) tensor decomposition of largest convolutional eigenvalues of RhoB in rectal cancer

A tensor is a multidimensional, multiway, or n -way array. A tensor of order one, order two, or order three is a vector, matrix, or volume, respectively. An n -th order or n -way tensor takes the form of an n -hypershaped. In many applications, data often represent a tensor with $n > 2$. In this study, a 3-way tensor is adopted to model the integration of the IHC images, color channels, and their sequences of maximum convolutional eigenvalues of RhoB expression in rectal cancer.

In general, the elements of a 3-way tensor, denoted as $\underline{\mathbf{T}}$ (the underline indicates a tensor), is equivalent to

$$t_{ijk} \approx \sum_{f=1}^F a_{if} b_{jf} c_{kf}, \quad (8)$$

where F is the number of factors, t_{ijk} are the elements of $\underline{\mathbf{T}}$, and a_{if} , $i = 1, \dots, I$, b_{jf} , $j = 1, \dots, J$, and c_{kf} , $k = 1, \dots, K$, are elements of three loading matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, respectively.

Alternatively, the tensor can be expressed as^{34,35}

$$\underline{\mathbf{T}} \approx \sum_{f=1}^F \mathbf{a}_f \otimes \mathbf{b}_f \otimes \mathbf{c}_f, \quad (9)$$

where $\otimes$ denotes the outer product, and $\mathbf{a}_f$, $\mathbf{b}_f$, and $\mathbf{c}_f$ are the f -column vectors of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$, respectively.

A solution for computing the loading matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be obtained using the PARAFAC decomposition model^{36,37}. The PARAFAC method can be considered as a generalization of bilinear PCA (principal component analysis). The use of the PARAFAC model has a certain advantage over other tensor-decomposition methods. The merit is that it provides unique solutions to the tensor decomposition by using the method of alternating least squares (ALS) to minimize the error of approximation with the model: $\mathbf{T} = \mathbf{a}(\mathbf{b} \otimes \mathbf{c})$, where the unfolded array $\mathbf{T} = I \times JK$.

For having more than one factor, let $\mathbf{b} \otimes \mathbf{c} = \mathbf{Y}$. Then the expression for $\mathbf{T}$ can be rewritten as

$$\mathbf{T} = \mathbf{A}\mathbf{Y}. \quad (10)$$

If $\mathbf{Y}$ is given by initializing $\mathbf{b}$ and $\mathbf{c}$, then $\mathbf{A}$ can be estimated as

$$\mathbf{A} = \mathbf{T}\mathbf{Y}^T(\mathbf{Y}\mathbf{Y}^T)^{-1}. \quad (11)$$

The estimates of $\mathbf{B}$ and $\mathbf{C}$ can be obtained using similar procedure for estimating $\mathbf{A}$. To compute $\mathbf{B}$, tensor $\underline{\mathbf{T}}$ is unfolded to matrix $\mathbf{T}$ of size $J \times IK$, and $\mathbf{Y}$ is determined from $\mathbf{A}$ and $\mathbf{C}$. Likewise, for the estimate of $\mathbf{C}$, $\mathbf{T}$ of size $K \times IJ$, and $\mathbf{Y}$ is determined from $\mathbf{A}$ and $\mathbf{B}$. The general procedure for estimating $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ can be outlined as follows³⁵.

PARAFAC Algorithm for a 3-Way Tensor:

1. Given F , initializing $\mathbf{B}$ and $\mathbf{C}$.
2. Compute $\mathbf{A}$ using $\underline{\mathbf{T}}$, $\mathbf{B}$, and $\mathbf{C}$.
3. Compute $\mathbf{B}$ and $\mathbf{C}$ likewise.
4. Stop if the solution converges or change between iterations is within tolerance. Otherwise, return to step 2.

Having described the general procedure of the PARAFAC decomposition of a 3-way tensor, The 3-way tensor for the cohort of patients having the survival rate of less than 5 years, denoted as $\underline{\mathbf{T}}_L$, can be modeled as

$$\underline{\mathbf{T}}_L = I_L \times C_L \times E_L, \quad (12)$$

where the subscript L stands for “less than 5 years”, I_L , C_L and E_L are the numbers of samples, image channels, and sequences of the maximum convolutional eigenvalues of the image channels, respectively, belonging to the cohort having the survival rate of less than 5 years.

Likewise, the 3-way tensor for the cohort of patients having the survival rate of more than 5 years, denoted as $\underline{\mathbf{T}}_M$, can be expressed as

$$\underline{\mathbf{T}}_M = I_M \times C_M \times E_M, \quad (13)$$

where the subscript M stands for “more than 5 years”, I_M , C_M and E_M are the numbers of samples, image channels, and sequences of the maximum convolutional eigenvalues of the image channels, respectively, belonging to the cohort having the survival rate of more than 5 years.

Figure 2 (B) shows the 3-way tensor model and its PARAFAC decomposition addressed in this paper for the multidimensional data analysis of RhoB expression in rectal cancer.

Results

To carry out the data analysis, all IHC biopsy samples were resized to 500×500 images. The sharpening kernel, pool size = 3×3 , stride = 1, and terminating value $n = 100$ were used for computing the sequence of largest convolutional eigenvalues of each IHC image channel by using the iterative image convolution procedure (Figure 2(A)). The computation resulted in sequences of 200 largest eigenvalues for each of the three color-image channels. However, it was observed that values of the largest eigenvalues after the first 12 points were very similar or the same for all images. Therefore, only largest-eigenvalue sequences of length = 12 (first 12 points) were used in the analysis. To construct previously described 3-way tensor models, the number of factors $F = 3$ was specified. The following results were obtained from the PARAFAC tensor decomposition.

Table 1 shows the means, standard deviations, p -values, and 95% confidence intervals of the three tensor-decomposition factors (F1, F2, and F3) of the two groups of patients.

Figures 3 and 4 show the 2D and 3D plots of F1, F2, and F3 obtained for the two groups of patients, respectively.

Figure 5 shows the distributions of F1, F2, and F3 with respect to the two groups of patients.

Discussion

The means and 95% confidence intervals of the 3 tensor-decomposition factors of the largest eigenvalues of IHC biopsy images of RhoB expression between the two groups of patients having survival rates < 5 years and > 5 years were found statistically significant (p -values $\ll .05$). The plots of the factors of the two groups of patients in both 2 and 3 dimensions as shown in Figures 3 and 4, where factors of the two groups are well-separated from each other, strongly suggest the predictive factor of RhoB in rectal cancer.

The predictive factor of RhoB is further confirmed by observing the distributions of tensor-decomposition factors F1, F2, and F3 along the corresponding biopsy samples of the two rectal-cancer patient groups, which are well-defined as shown in Figures 5(a) and (b) for F1 and F2, respectively. Although the plots of F3 factor for both groups partially overlap (Figure 5(c)), their means are distinct (Table 1).

While manual analysis of RhoB expression in IHC biopsy samples of rectal-cancer patients carried out by pathologists could not discover the predictive power of this protein coding gene (our group's unpublished data), the tensor decomposition of largest convolutional eigenvalues of RhoB expression objectively suggests that its measurement is associated with patients' survival. Due to highly intricate color and texture of protein expression on images, it is extremely difficult if not possible for the human eye to discern many parts of a sample that are spatially arranged in very complicated or delicate structures.

To allow timely and efficient processes of biomarker discovery in cancer research, information about the spatial cellular composition and heterogeneity of tissues need to be accurately identified. This is because insightful knowledge of cell subpopulations and the tumor microenvironment through TMAs and IHC is important to elucidate the biology and development of cancer³⁸. The conventional assessment of protein expression on IHC imaging in pathology is not only prone to qualitative and quantitative errors and disagreements, but also difficult to make the results reproducible, particularly for a large amount of data. Limitation and imprecision of manual scoring of protein expression on TMAs have already been well aware and several methods for automated image analysis developed^{39–41}. The combination of advanced AI and computational data analysis methods presented in this study has enabled the discovery by objectively showing the predictive factor of RhoB expression. First, the presented approach provides a rigorous and effective image feature extraction of the protein expression by using a deep learning-like procedure for computing “deep” convolutional eigenvalues of the image. The largest eigenvalue is equivalent to the first principle component that accounts for the largest variance of the data. By considering the largest eigenvalues determined after a series of image convolution, both spatial transformation and temporal information of complex color distribution and texture can be captured in the process. Second, multidimensional data is then modeled as a multiway array and tensor decomposition used to analyze the fusion of latent information. Finally, analysis results are then represented with compact informative properties as components that can precisely differentiate subtle differences in the protein expression between patient groups.

Technical advantages of the data processing of the presented methodology are of two aspects. First, because the AI-based image processing is capable of handling samples of large sizes by downsizing to produce output images that are much smaller than the input, the method is thus scalable. Second, because sequences of the largest eigenvalues of convolved images can be truncated at some much shorter length due to the quick convergence of the values, the computation of decomposition factors can be more efficient in both time and storage, reducing computational complexity.

Applications of state-of-the-art AI and data science in precision or personalized medicine optimistically hold promises for medical revolution aiming to achieve effective treatments and cures of human diseases that are leading causes of death, such as cancer, cardiovascular diseases, neurodegenerative disorders, and rare genetic conditions^{42–46}, all of which depend on the accuracy of biomarker assessment. In addition to AI, in fact, tensor decomposition or factorization has been realized as a potential method for precision medicine in studying multiple modalities of biomedical data to mine useful information hidden from the original data⁴⁷. Similarly, the method of coupled matrix and tensor decomposition was utilized as an information fusion scheme for discovering diagnostic biomarkers of schizophrenia using EEG, functional MRI, and structural MRI data⁴⁸.

In summary, RhoB expression in a randomized preoperative clinical trial of rectal-cancer patients with long follow-up data has been addressed with the use of the unique combination of AI-based and multidimensional data analysis methods. This automated identification of the predictive power of RhoB expression in rectal cancer is promising for contributing to leveraging the accuracy and efficiency of biomarker discovery and facilitating novel developments of clinical trials.

Conclusions

Methods for extracting deep image features of RhoB expression in rectal-cancer patients and tensor decomposition of these features known as largest convolutional eigenvalues have been presented and discussed in the foregoing sections. The results obtained from this study demonstrates the potential power of RhoB as a predictive factor for rectal cancer. Furthermore, the methods presented herein is not limited to the pathological image analysis of the protein expression in rectal cancer but can be utilized for biomarker discovery in other types of cancer using microscope-based screening data.

The increasing availability of “big data” in biomedicine offers tremendous potentials for the development and applications of personalized medicine aiming at achieving optimal treatment and effective care of patients^{49–51}. However, there are computational challenges for the use of big data analytics, such as the preprocessing and analysis of whole slide imaging in digital pathology, which need to be overcome before personalized medicine can be put into clinical settings. The scalability and computational efficiency of the methods addressed in this study appear to be a useful tool contributing to extracting meaningful insights from very large amounts of biomedical data.

Software availability

MATLAB codes and data used in this paper are available at the first and corresponding author’s personal homepage: <https://sites.google.com/view/tuan-d-pham/codes> under the title “Tensor decomposition of RhoB eigenvalues”.

References

1. Siegel, R., Miller, K., Jemal, A. Cancer statistics, 2020. *CA Cancer J Clin.* **70**, 7-30 (2020).
2. Ferlay, J., Soerjomataram, I., Dikshit, R., et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* **136**, E359-E386 (2015).
3. Bosset, J.F., Calais, G., Mineur, L., et al. Fluorouracil-based adjuvant chemotherapy after preoperative chemoradiotherapy in rectal cancer: Long-term results of the EORTC 22921 randomised study. *Lancet Oncol* **15**, 184-190 (2014).
4. Naffouje, S., Sabesan, A., Powers, B.D., et al. Patient risk subgroups predict benefit of adjuvant chemotherapy in stage II rectal cancer patients following neoadjuvant chemoradiation and total mesorectal excision. *Clin Colorectal Cancer* **20**, e155-e164 (2021).
5. Ridley, A.J. Rho family proteins: coordinating cell responses. *Trends Cell Biol.* **11**, 471-417 (2001).
6. Ridley, A.J. Rho proteins: linking signaling with membrane trafficking. *Traffic* **2**, 303-310 (2001).
7. Prendergast, G.C. Actin’ up: RhoB in cancer and apoptosis. *Nat Rev Cancer* **1**, 162-168 (2001).
8. Luis-Ravelo, D., Anton, I., Zanduetta, C., et al. RHOB influences lung adenocarcinoma metastasis and resistance in a host-sensitive manner. *Mol Oncol.* **8** 196-206 (2014).
9. Ma, Y., Gong, Y., Cheng, Z., et al. Critical functions of RhoB in support of glioblastoma tumorigenesis. *Neuro Oncol.* **17**, 516-525 (2015).
10. Chen, W., Niu, S., Ma, X., et al. RhoB acts as a tumor suppressor that inhibits malignancy of clear cell renal cell carcinoma. *PLoS ONE* **11** e0157599 (2016),
11. Ju, J.A., Gilkes, D.M. RhoB: Team oncogene or team tumor suppressor?. *Genes* **9**, 67 (2018)
12. Vega, F.M., Ridley A.J. The RhoB small GTPase in physiology and disease. *Small GTPases* **9**, 384-393 (2018).
13. Ju, J.A., Godet, I., DiGiacomo, J.W., Gilkes, D.M. RhoB is regulated by hypoxia and modulates metastasis in breast cancer. *Cancer Reports* **3**, e1164 (2020).
14. Arsic, N., Ho-Pun-Cheung, A., Lopez-Crapez, E., et al. The p53 isoform delta133p53 β regulates cancer cell apoptosis in a RhoB-dependent manner. *PLoS ONE* **12**, e0172125 (2017).
15. Liu, N., Cui, W., Jiang, X., et al. The critical role of dysregulated RhoB signaling pathway in radioresistance of colorectal cancer. *Int J Radiation Oncol Biol Phys* **104**, 1153-1164 (2019).
16. Giordani, P., Rocci, R., Bove, G. Factor uniqueness of the structural Parafac model. *Psychometrika* **85**, 555-574 (2020).
17. von Chamier, L., Laine, R.F., Jukkala, J. et al. Democratizing deep learning for microscopy with ZeroCostDL4Mic. *Nat Commun* **12**, 2276 (2021).
18. Chen, M., Ludtke, S.J. Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. *Nat Methods* **18**, 930–936 (2021).
19. Speiser, A., Müller, L.R., Hoess, P., et al. Deep learning enables fast and dense single-molecule localization with high accuracy. *Nat Methods* **18**, 1082-1090 (2021).
20. Smilde, A., Bro, R., Geladi, P. *Multi-way Analysis: Applications in the Chemical Sciences*. Wiley, Chichester (2004).
21. Kolda, T.G., Bader, B.W. Tensor decompositions and applications. *SIAM J. Matrix Anal. Appl.* **51**, 455-500 (2009).
22. Luo, Y., Ahmad, F.S., Shah, S. J. Tensor factorization for precision medicine in heart failure with preserved ejection fraction. *J Cardiovasc Transl Res.* **10**, 305-312 (2017)

23. Pham, T.D., Yan, H. Tensor decomposition of gait dynamics in Parkinson's disease, *IEEE. Trans. Biomed. Eng.* **65**, 1820-1827 (2018).
24. Zhao, J., Zhang, Y., Schlueter, D.J., et al. Detecting time-evolving phenotypic topics via tensor factorization on electronic health records: Cardiovascular disease case study. *J. Biomed. Inform* **98**, 103270 (2019).
25. Wang, R., Li, S., Cheng, L. et al. Predicting associations among drugs, targets and diseases by tensor decomposition for drug repositioning. *BMC Bioinformatics* **20**, 628 (2019).
26. Cirillo, M.D., Mirdell, R., Sjoberg, F., Pham, T.D. Tensor decomposition for color image segmentation of burn wounds. *Sci. Rep.* **9**, 329 (2019).
27. Korevaar, H., Metcalf, C.J., Grenfell, B.T. Tensor decomposition for infectious disease incidence data. *Methods Ecol Evol.* **11**, 1690-1700 (2020).
28. Nesaragi, N., Patidar, S., Aggarwal, V. Tensor learning of pointwise mutual information from EHR data for early prediction of sepsis. *Comput. Biol. Med.* **134**, 104430 (2021).
29. Ng, K.L., Taguchi, Y.H. Identification of miRNA signatures for kidney renal clear cell carcinoma using the tensor-decomposition method. *Sci Rep* **10**, 15149 (2020).
30. Leistico, J.R., Saini, P., Futtner, C.R., et al. Epigenomic tensor predicts disease subtypes and reveals constrained tumor evolution. *Cell Rep.* **34**, 108927 (2021).
31. Swedish Rectal Cancer Trial. Improved survival with preoperative radiotherapy in resectable rectal cancer. *N Engl J Med.* **8**, 980-987 (1997).
32. Pham, T.D. Convolutional fuzzy recurrence eigenvalues. *EPL*, accepted manuscript available online 23 June 2021. DOI: <https://doi.org/10.1209/0295-5075/ac0df8>
33. Gonzalez, R.C., Woods, R.E. *Digital Image Processing*, 4th edition (Pearson, 2018)
34. Burdick, D.S. An introduction to tensor products with applications to multiway data analysis. *Chemom. Intell. Lab. Syst.* **28**, 229-237 (1995).
35. Bro, R. PARAFAC. Tutorial and applications. *Chemom. Intell. Lab. Syst.* **38**, 149-171 (1997).
36. Harshman, R.A. Foundations of the PARAFAC procedure: Model and conditions for an 'explanatory' multi-mode factor analysis, *UCLA Working Papers in Phonetics* **16** 1-84 (1970).
37. Carroll, J.D., Chang, J. Analysis of individual differences in multidimensional scaling via an *N*-way generalization of "Eckart-Young" decomposition. *Psychometrika* **35**, 283-319 (1970).
38. Blom, S., Paavolainen, L., Bychkov, D., et al. Systems pathology by multiplexed immunohistochemistry and whole-slide digital image analysis. *Sci Rep* **7**, 15580 (2017).
39. Ali, H., Irwin, M., Morris, L. et al. Astronomical algorithms for automated analysis of tissue protein expression in breast cancer. *Br J Cancer* **108**, 602-612 (2013).
40. Kumar, A. Rao, A., Bhavani, S., et al. Automated analysis of cancer location biomarkers, *PNAS* **111**, 18249-18254 (2014).
41. Acs, B., Rantalainen, M., Hartman, J. Artificial intelligence as the next step towards precision pathology. *J Intern Med* **288**, 62-81 (2020).
42. Troyanskaya, O., Trajanoski, Z., Carpenter, A. et al. Artificial intelligence and cancer. *Nat Cancer* **1**, 149-152 (2020).
43. Uddin, M., Wang, Y., Woodbury-Smith, M. Artificial intelligence for precision medicine in neurodevelopmental disorders. *npj Digit. Med.* **2**, 112 (2019).
44. Kermany, D.S., Goldbaum, M., Ca, W., et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **172**, 1122-1131.e9 (2018).
45. Subramanian, M., Wojtusciszyn, A., Favre, L., et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. *J Transl Med* **18**, 472 (2020).
46. Johnson, K.B., Wei, W.Q., Weeraratne, D., et al. Precision medicine, AI, and the future of personalized health care. *Clin Transl Sci.* **14**, 86-93 (2021).
47. Luo, Y., Wang, F., Szolovits, P. Tensor factorization toward precision medicine. *Brief. Bioinformatics* **18**, 511-514 (2017).
48. Evrim, A., Carla, S., Yuri, L.S., et al. Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data. *Front. Neurosci.* **13**, 416 (2019).

49. Costa, F.F. Big data in biomedicine. *Drug Discov Today*. **19**, 433-440 (2014).
50. Bender, E. Big data in biomedicine. *Nature* **527**, S1 (2015).
51. Luo, J., Wu, M., Gopukumar, D., Zhao, Y. Big data application in biomedical research and health care: A literature review. *Biomed. Inform. Insights* **8**, 1–10 (2016),
52. Kumar, N., Gupta, R., Gupta, S. Whole slide imaging (WSI) in pathology: Current perspectives and future directions. *J Digit Imaging* **33**, 1034-1040 (2020).

Acknowledgments

T.D.P. acknowledged institutional support for the PMU Center for Artificial Intelligence. X.F.S. acknowledged support from the Swedish Cancer Foundation, Oncology Clinics, and LiU Cancer.

Authors contribution

T.D.P conceptualized and designed the mathematical methods, implemented the methods, and carried out the computer experiments. T.D.P. wrote the manuscript. V.R. assisted T.D.P. in the computer implementation and participated in the numerical analysis. B.L. and C.W.F. performed the TMA scanned image arrangement and splicing. N.L. carried out the IHC experiment, X.F.S. coordinated the work with N.L., B.L., and C.W.F. on pathological aspects of rectal cancer. All authors reviewed the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

Table 1. Means, standard deviations (SD), 95% confidence intervals (95% CI), and *p*-values of distributions of three tensor-decomposition (TD) factors of largest convolutional eigenvalues of biopsy samples of two groups of patients.

TD Factor	Mean $\pm$ SD	95% CI	<i>p</i> -value
< 5 years			
F1	734.8288 $\pm$ 48.7095	(720.1949, 749.4628)	9.2419e-54
F2	445.8856 $\pm$ 32.2876	(436.1853, 455.5859)	4.4357e-52
F3	119.1925 $\pm$ 15.6983	(114.4762, 123.9088)	9.2989e-41
> 5 years			
F1	1.1030e+03 $\pm$ 59.9351	(1.0892e+03, 1.1168e+03)	1.2657e-95
F2	-203.6226 $\pm$ 42.6095	(-213.4261, -193.8190)	6.4535e-53
F3	53.2512 $\pm$ 28.7499	(46.6364, 59.8659)	8.5563e-26

Figure 1. IHC biopsy samples

Figure 2. Graphical depiction of (A) procedure for computing sequence of largest convolutional eigenvalues of an image, where “C&R” denotes “Convolution & ReLU”; and (B) PARAFAC decomposition of 3-way tensor used in this study.

(a)

(b)

(c)

Figure 3. Plots of pairwise distributions of tensor decomposition factors of largest convolutional eigenvalues of RhoB expression in biopsy of two patient groups of survival.

Figure 4. Three-dimensional plot of distributions of tensor decomposition factors of largest convolutional eigenvalues of RhoB expression in biopsy of two patient groups of survival.

Figure 5. Distributions of tensor-decomposition factors of largest convolutional eigenvalues of RhoB expression in biopsy of two patient groups of survival.