Luigi: Large-scale histopathological image retrieval system using deep texture representations

Corresponding author: Shumpei Ishikawa; 113-8510, sish.gpat@mri.tmd.ac.jp, +81-3-5803-4817

Authors: Daisuke Komura¶[1], Keisuke Fukuta¶[1,2], Ken Tominaga[1,3], Akihiro Kawabe[1,4], Hirotomo Koda[1,4], Ryohei Suzuki[1,5], Hiroki Konishi[1,2], Toshikazu Umezaki[1], Tatsuya Harada[2], Shumpei Ishikawa[1]

[1]Department of Genomics Pathology, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

[2]Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan

[3]Graduate School of Interdisciplinary Information Studies, The University of Tokyo, Tokyo, Japan

[4]Department of Pathology and Diagnostic Pathology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[5]Department of Physics, Graduate School of Science, The University of Tokyo

¶Contributed equally.

Abstract

Background:

As a large number of digital histopathological images have been accumulated, there is a growing demand of content-based image retrieval (CBIR) in pathology for educational, diagnostic, or research purposes. However, no CBIR systems in digital pathology are publicly available.

Results:

We developed a web application, the Luigi system, which retrieves similar histopathological images from various cancer cases. Using deep texture representations computed with a pre-trained convolutional neural network as an image feature in conjunction with an approximate nearest neighbor search method, the Luigi system provides fast and accurate results for any type of tissue or cell without the need for further training. In addition, users can easily submit query images of an appropriate scale into the Luigi system and view the retrieved results using our smartphone application. The cases stored in the Luigi database are obtained from The Cancer Genome Atlas with rich clinical, pathological, and molecular information. We tested the Luigi system and the smartphone application by querying typical cancerous regions from four cancer types, and confirmed successful retrieval of relevant images with both applications.

Conclusions:

The Luigi system will help students, pathologists, and researchers easily retrieve histopathological images of various cancers similar to those of the query image.

Luigi is freely available at https://luigi-pathology.com/.

Keywords

Machine learning, Image analysis, Information retrieval, Neural networks, Web services

Background

Recently, computational image analysis has been applied to the ever increasing volume of digital histopathological and medical images for many purposes[1,2]. Content-based image retrieval (CBIR) is one such application that facilitates retrieval of the most relevant images to a query image from an image database. In digital pathology, CBIR systems are useful in many situations, particularly diagnosis, education, and research. For example, CBIR systems can be used for educational purposes by students and beginner pathologists to retrieve relevant cases or histopathological images of tissues. In addition, such systems are also helpful to professional pathologists, particularly with the diagnosis of rare cases.

Several CBIR systems to retrieve similar histopathological images have been proposed[3,4]; however, to the best of our knowledge, no publicly available system has yet been released. Therefore, we developed Luigi, a novel web-based CBIR system, which allows users to easily retrieve similar histopathological images of cancer cases with various useful annotations.

The Luigi system calculates image relevance based on representations from a deep convolutional

neural network (CNN), which has achieved great success in a wide range of visual tasks. Well-trained deep neural networks provide visual features that perform well in a wide range of computer vision applications, including medical image analysis[5–8]. Although some CBIR applications have attempted to employ deep learning, such applications require label information, which is difficult to obtain in the field of pathology and is not scalable to large and various datasets. Based on the observation that pathological images, particularly of various cancer cases, potentially have textural structures and that spatial autocorrelation is often used in texture analysis, rather than using features from a CNN directly, we adopted the autocorrelation of the feature maps of a CNN pre-trained on ImageNet, which is referred to as deep texture representations[9]. This approach works remarkably well with pathological images without the necessity of additional learning processes for specific datasets. Therefore, the Luigi system is easily applicable to various tissues and scalable to a large number of histopathological images. A smartphone application was also developed that allows users to easily submit cases and view the retrieved results without the need for a dedicated system. Moreover, as all cancer cases in Luigi are retrieved from The Cancer Genome Atlas (TCGA)[10], which is a collection of more than 10000 whole slide images (WSIs) of 32 cancer types with rich clinical, pathological, and molecular information, pathologists and researchers can also obtain further insight into the molecular basis of the cancer of interest. The effectiveness of the Luigi CBIR system was demonstrated by querying typical cancerous regions.

Implementation

The Luigi system was developed with a Python back end and JavaScript front end; thus, it can be accessed by any modern JavaScript-enabled web browser with a Linux, Mac OS X, or Windows operating system. An overview of the Luigi system is shown in Figure 1.

The Luigi system can be used by accessing its URL via any modern web browser. Users can search relevant images relevant to a query image from the database using intuitive operations. Users can also view the cBioPortal[11] website to view clinical, pathological, and molecular information relevant to the retrieved case. Although users do not need to register accounts to use the Luigi system, if an account is registered, the system saves a history of all uploaded images.

The Luigi database contains WSI information for each case; however, the query and retrieved images handled by the system are patches extracted from WSIs because in most situations, users submit small regions of a whole slide with specific features[4].

Details about the system are described in the following sections.

Retrieval system

A CBIR system requires quantitative metrics that represent image “relevance.” Therefore, effective features must be extracted to compute the similarity between histopathological images. For example, the similarity of histopathological images is often considered within the context of cell structure, tissue texture, or shape of the nucleus. Many researchers have applied CBIR to medical domains

by employing hand-crafted features or classical texture features based on such information. In contrast, the Luigi system uses deep texture representations computed using a pre-trained CNN for feature extraction of images, an approach motivated by recent studies of neural style transfer[9]. The deep texture representations are computed using a correlation matrix of feature maps in a CNN layer. Here the “conv3 1” layer of VGG-16[12] was employed.

$$G_{ij} = \sum_k F_{ik} F_{jk}$$

In the above equation, G, F, and k represent the correlation matrix, the feature maps of the CNN, and the position in the images, respectively. This manipulation, which is also called bilinear pooling, produces orderless image representations by aggregating the averages of the location-wise outer products of the feature maps. Note that this process discards some spatial information; therefore, the representations have spatial invariance, which appears to be necessary for pathological images because the similarity between two images should not be affected by the relative location of cells within the image. Because the representations are too high dimensional for practical application, the Compact Bilinear Pooling (CBP)[13] was adopted to approximate the original representations of 65536 dimensions with lower dimensions.

CBP is an approximation method of bilinear pooling using two random matrices and has good properties without the need for pre-training, unlike other dimension reduction methods, such as principal component analysis. The final representations for each image have 1024 dimensions, which demonstrated good performance when evaluating histopathological images.

When a query image is input into the Luigi system, deep texture representations of the image are created and rotated by 90°, 180°, and 270° to impose rotational invariance and compress the representations of each image using CBP. The system then searches for the nearest images in the database using the cosine similarity of the extracted compressed features between the query and database images.

To enable high-speed searching, an approximate nearest neighbor search method with the randomized kd-tree algorithm was adopted. Overall operations, including feature extraction and nearest search, require less than 2 s per image, which is sufficient for practical applications.

Dataset and pre-processing

An image database was created from WSIs of all cancer types in the TCGA. The SVS files of the diagnostic WSIs that include DX in the file name were downloaded from the National Cancer Institute GDC legacy archive[14]. WSIs without sufficient information, such as magnification level, were discarded. Finally, the database contained WSIs of a total of 9662 patients and 32 cancer types (Table 1). First, image patches of 256 x 256 pixels are extracted from approximately 50% of the tissue area for each WSI. The tissue area is detected using the Otsu algorithm[15]. Typically, pathologists evaluate histopathological images at multiple magnification levels; thus, patches were extracted at two unique scales (10x and 20x). Second, deep texture representations were calculated for each image patch. For memory efficiency, similar patches in the same slide were reduced. k-

means clustering (k = 500) was adopted and only the nearest patch from each centroid was used.

Table 1. Cancer types included in the database. (placed at the end of this file)

User interface of the Luigi system

Figure 2 shows the interface of the Luigi system. Users upload a histopathological image file by clicking the “Upload File” button or directly dragging and dropping the image file. The uploaded image is resized to 256 x 256 pixels, which is the same as the size of images in the database. Because this resizing process could change the texture, it is important to upload images with appropriate magnification levels to retrieve good results. Alternatively, images randomly selected from the TCGA can be input to a query by choosing the cancer types and clicking on one of the images in the lower left on the screen. Then the Luigi system will retrieve the most relevant patch in each WSI, rank each according to the cosine similarity, and initially display the top 48 relevant adrenal gland tumor cases at 10x magnification. Cancer types and magnification can be changed by clicking the names/magnifications on top of the results retrieved column. When a retrieved image is clicked, the enlarged view will appear. The enlarged image can be further input to a query by clicking “Search this” button. When the “Jump to cBioPortal” button below the enlarged image or the TCGA case ID (TCGA-XX-XXX) above each retrieved image is clicked, a new window will open and the summary of the selected case will be displayed on the cBioPortal website, which enables

the user to investigate the pathology report as well as clinical and molecular information of the case.

Smartphone application

A smartphone application was also developed that allows users to directly submit appropriately scaled query images into the Luigi web application and view the retrieved images on a smartphone. Figure 3 shows the interface of the Luigi smartphone application. First, users select a reference image from the list. By swiping the screen, the magnification of the reference image is changed. Then users can take a photo of their own cases by tapping anywhere in the screen. Users can easily adjust the size of a query image using the reference image. Immediately after capturing a query image, the top 48 relevant cases of adrenal gland tumors at the same magnification level of the selected reference image will be displayed initially. Cancer types can be changed by swiping the screen. Users can also access the cBioPortal website by selecting a retrieved case. Currently, only an iOS version of the smartphone application is available, but an Android version will be developed in the future.

Results

To determine if the Luigi system can accurately retrieve relevant images and evaluate how query images captured by the smartphone application affect the results, as compared to the original images, the effectiveness of the Luigi system was assessed by querying some typical cancerous

regions from TCGA images. Queries and the corresponding image database used in this experiment are summarized in Table 2. For each query, the most relevant images in each WSI are retrieved and ranked according to the cosine similarity and then the averaged precision of the top 20 images for three queries in each query cancer type is calculated to measure the retrieval accuracy. For the evaluation of query images captured by the smartphone application, the original images were displayed on 24-inch 4K LCD monitor and photos were captured three times using an iPhone X camera. Then, the averaged precision was compared to that of the original query image. The query images (Figure 4), which show the typical appearance of each query cancer type, were selected by a pathologist and the retrieved images were manually evaluated by a second pathologist. Note that because these WSIs contain both cancerous and non-cancerous tissues, including normal alveolar epithelium, muscle, cartilage, red blood cells, and leukocytes, patches containing only non-cancerous regions were regarded as incorrect even if they belong to the WSIs of the correct cancer type.

Table 2. Queries and the corresponding image database used in the experiment.

Query cancer type	Abbreviation	Cancer types in the image database
Cancerous regions of lung adenocarcinoma	LUAD	Lung adenocarcinoma, lung squamous cell carcinoma and mesothelioma
Cancerous regions of lung squamous cell carcinoma	LUSC	Lung adenocarcinoma, lung squamous cell carcinoma and mesothelioma
Cancerous regions of chromophobe renal cell carcinoma	KCP	Chromophobe renal cell carcinoma, clear renal cell carcinoma and papillary renal cell carcinoma
Cancerous regions of renal	KRCC	Chromophobe renal cell carcinoma, clear

clear cell carcinoma		renal cell carcinoma and papillary renal cell carcinoma

Figure 5 presents the precision obtained by Luigi's CBP based on VGG-16, GoogLeNet (inception_4b layer) [16] and ResNet50 (res2 layer) [17], and rotation invariant local binary pattern (LBP)[18], which is one of the common texture feature descriptors in image analysis. The figure shows that the Luigi CBP based on VGG-16 outperformed the LBP for all query types except for queries two and three in KRCC, where both methods show equally perfect results, and query three in LUSC, where LBP shows slightly better performance. Additionally, most of the incorrect images appeared similar to the query image even to experienced pathologists, as shown in Figures 6 and 7, possibly because the classification of a cancer type is not always strict and various tumor tissues often contain multiple types of cancer, which frequently leads to interobserver disagreement[19]. Additionally, different categories of cancers often have similar appearances in some parts of the tumors, which could be indistinguishable even by trained pathologists. For example, renal clear renal cell carcinoma looks similar to chromophobe renal cell carcinoma, as shown in Figure 6(A) , and lung adenocarcinoma cells look similar to lung squamous cell carcinoma in some patches, as shown in Figure 7(B). The evaluation criteria used in this experiment seem to slightly decrease inevitably by such biological factors. Nevertheless, these results demonstrate that the Luigi system successfully retrieved relevant images regardless of the cancer type, which is the purpose of the system. The experimental results also showed that the retrieval accuracies were comparable among

VGG-16, GoogLeNet and ResNet50. However, the improvement of the retrieval accuracy of the VGG-16 based method over the LBP method was more stable than that of GoogLeNet and ResNet50 across the datasets. Thus, we adopted VGG-16 in the Luigi system.

Figure 8 shows a comparison of the precision of the Luigi CBP between the original query images and corresponding photos captured by the smartphone version of Luigi. The precision of the photos captured by the smartphone application was comparable to that of the original query images in KRCC and LUAD. For LUSC and KCP, the precision of the photos captured by the smartphone application was worse than that of the original image. It is possibly because the typical appearance of these cancers often looks similar to cancer cells of other types as described in the previous experiment, and the difference between them could be intolerant to noise introduced by recapturing of images with a smartphone camera. Reducing aliasing of images recaptured from an LCD monitor [20] or capturing pictures obtained with a microscope directly using smartphone adapters[21] could improve performance. Notably, in this experiment, out-of-focus pictures significantly reduced precision. Thus, photos should be large and captured from a distance farther than the shortest focal length.

Conclusions

We developed a novel web-based CBIR application, the Luigi system, which enables pathologists, students, and researchers to retrieve cancer cases with histopathological images similar to those of

the query image. Experimental results showed that the Luigi system successfully retrieved relevant images. The features of Luigi system are as follows:

- To the best of our knowledge, this is the first publicly available CBIR system with rich clinical and genomic information in the pathology field.
- The Luigi image retrieval system uses deep texture representations to calculate the relevance metrics between pathological images.
- This enables the Luigi system to retrieve relevant images regardless of spatial information, e.g., the relative location of cells within the image.
- The Luigi system does not require extra learning; thus, it is scalable to a large number of histopathological images and any type of tissue or cell can be searched without training.
- The smartphone application allows users to easily submit query images without the need for special equipment and view the retrieved results on a smartphone.

We plan to evaluate the retrieval accuracy of the deep texture features from the CNN model, which is pre-trained in unsupervised way or fine-tuned using histopathological images. Each method has several issues to be resolved to improve the performance, but these are beyond the scope of this paper. For example, there are several options and hyperparameters for pre-training CNN using unsupervised learning such as learning methods (e.g. Restricted Boltzmann Machine, autoencoder, or Generative Adversarial Network), network structures, and learning rates. Fine-tuning could help improve the performance if histopathological images can be prepared with appropriate labels of

various tissue types with an unbiased way. Although these ideas have not yet been implemented, we believe that the performance of the Luigi system is both satisfactory and practical.

Availability and requirements

Project name: Luigi

Webserver: https://luigi-pathology.com/

iPhone application: available on the Apple iTunes store as “Luigi”

Operating systems: Platform independent (Webserver), iOS 10.0 or later (iPhone application)

Programming language: Python and Javascript (Webserver), Swift (iPhone application)

Other requirements: Firefox, Chrome, IE, or Safari. JavaScript must be enabled (Webserver).

License: None

Any restrictions to use by non-academics: Luigi cannot be used for commercial use.

Abbreviations

CBIR, Content-based image retrieval; CNN, convolutional neural network; TCGA, The Cancer Genome Atlas; WSI, whole slide image; CBP, Compact Bilinear Pooling.

Declarations

Ethics approval

All image data related to human subjects used for this study is de-identified and publicly available from The Cancer Genome Atlas project. Therefore, this research is not classified as a human subject research and no Institutional Review Board approval is required.

Consent for publication

Not applicable

Availability of data and materials

The image datasets analyzed in this study are available in the National Cancer Institute GDC Data Portal, https://portal.gdc.cancer.gov/

Funding

This work was supported by the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (A) [16H02481 to S.I.], by the Practical Research for Innovative Cancer Control from Japan Agency for Medical Research and development (S.I.), and by the Joint Usage/Research Program of Medical Research Institute, Tokyo Medical and Dental University (D.K., S.I.).

Competing interests

There are no competing interests.

References

1. Yu K-H, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016;7:12474.

2. Kothari S, Phan JH, Stokes TH, Wang MD. Pathology imaging informatics for quantitative analysis of whole-slide images. J Am Med Inform Assoc. 2013;20:1099–108.

3. Zhang X, Liu W, Dundar M, Badve S, Zhang S. Towards Large-Scale Histopathological Image Analysis: Hashing-Based Image Retrieval. IEEE Trans Med Imaging. 2015;34:496–506.

4. Qi X, Wang D, Rodero I, Diaz-Montes J, Gensure RH, Xing F, et al. Content-based histopathology image retrieval using CometCloud. BMC Bioinformatics. 2014;15:287.

5. Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A Deep Convolutional Neural Network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomputing. 2016;191:214–23.

6. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep. 2016;6:26286.

7. Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. ArXiv150407947 Cs [Internet]. 2015 [cited 2017 May 18]; Available from: http://arxiv.org/abs/1504.07947

8. Kleesiek J, Urban G, Hubert A, Schwarz D, Maier-Hein K, Bendszus M, et al. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping. NeuroImage. 2016;129:460–9.

9. Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. Proc IEEE Conf Comput Vis Pattern Recognit. 2016. p. 2414–2423.

10. Weinstein JN, Collisson EA, Mills GB, Shaw KRM, Ozenberger BA, Ellrott K, et al. The cancer genome atlas pan-cancer analysis project. Nat Genet. 2013;45:1113–1120.

11. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. AACR; 2012.

12. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. ArXiv Prepr ArXiv14091556. 2014;

13. Gao Y, Beijbom O, Zhang N, Darrell T. Compact Bilinear Pooling. ArXiv151106062 Cs [Internet]. 2015; Available from: http://arxiv.org/abs/1511.06062

14. National Cancer Institute GDC Data Portal [Internet]. [cited 2017 May 29]. Available from: https://portal.gdc.cancer.gov/

15. Otsu N. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst Man Cybern. 1979;9:62–6.

16. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going Deeper with Convolutions. ArXiv14094842 Cs [Internet]. 2014 [cited 2018 Feb 14]; Available from: http://arxiv.org/abs/1409.4842

17. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. ArXiv151203385 Cs [Internet]. 2015 [cited 2018 Feb 14]; Available from: http://arxiv.org/abs/1512.03385

18. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns - IEEE Journals & Magazine [Internet]. [cited 2018 Feb 14]. Available from: http://ieeexplore.ieee.org/document/1017623/

19. Longacre TA, Ennis M, Quenneville LA, Bane AL, Bleiweiss IJ, Carter BA, et al. Interobserver agreement and reproducibility in classification of invasive breast carcinoma: an NCI breast cancer family registry study. Mod Pathol. 2005;19:195–207.

20. Muammar H, Dragotti PL. An investigation into aliasing in images recaptured from an LCD monitor using a digital camera. 2013 IEEE Int Conf Acoust Speech Signal Process. 2013. p. 2242–6.

21. Roy S, Pantanowitz L, Amin M, Seethala RR, Ishtiaque A, Yousem SA, et al. Smartphone adapters for digital photomicrography. J Pathol Inform. 2014;5:24.

Figures

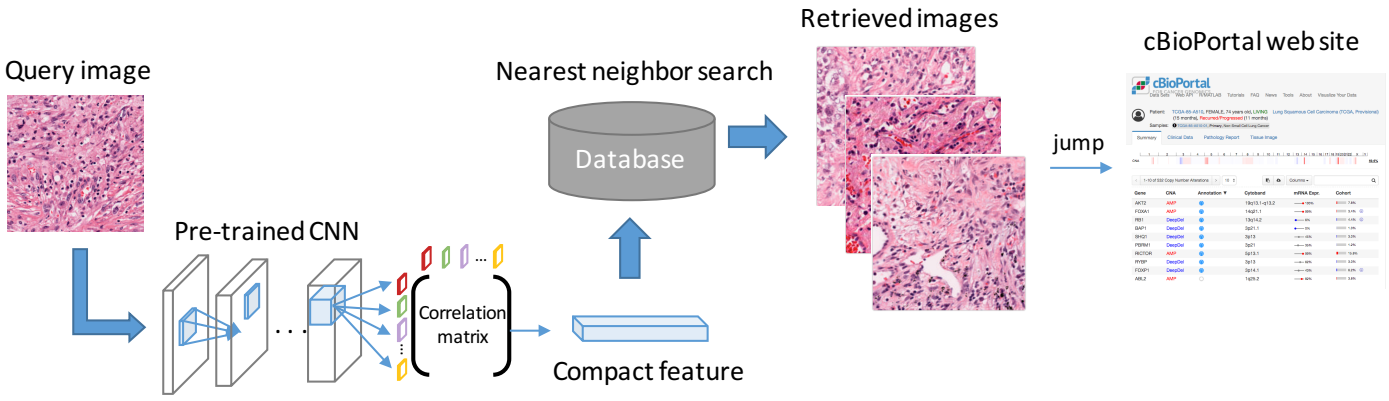


Figure 1. Overview of the Luigi system. After submitting the query image, deep texture representations are calculated and compressed by compact bilinear pooling, and the 48 most similar images are retrieved. Users can view the retrieved cases on the cBioPortal website by clicking the link to the selected cases.

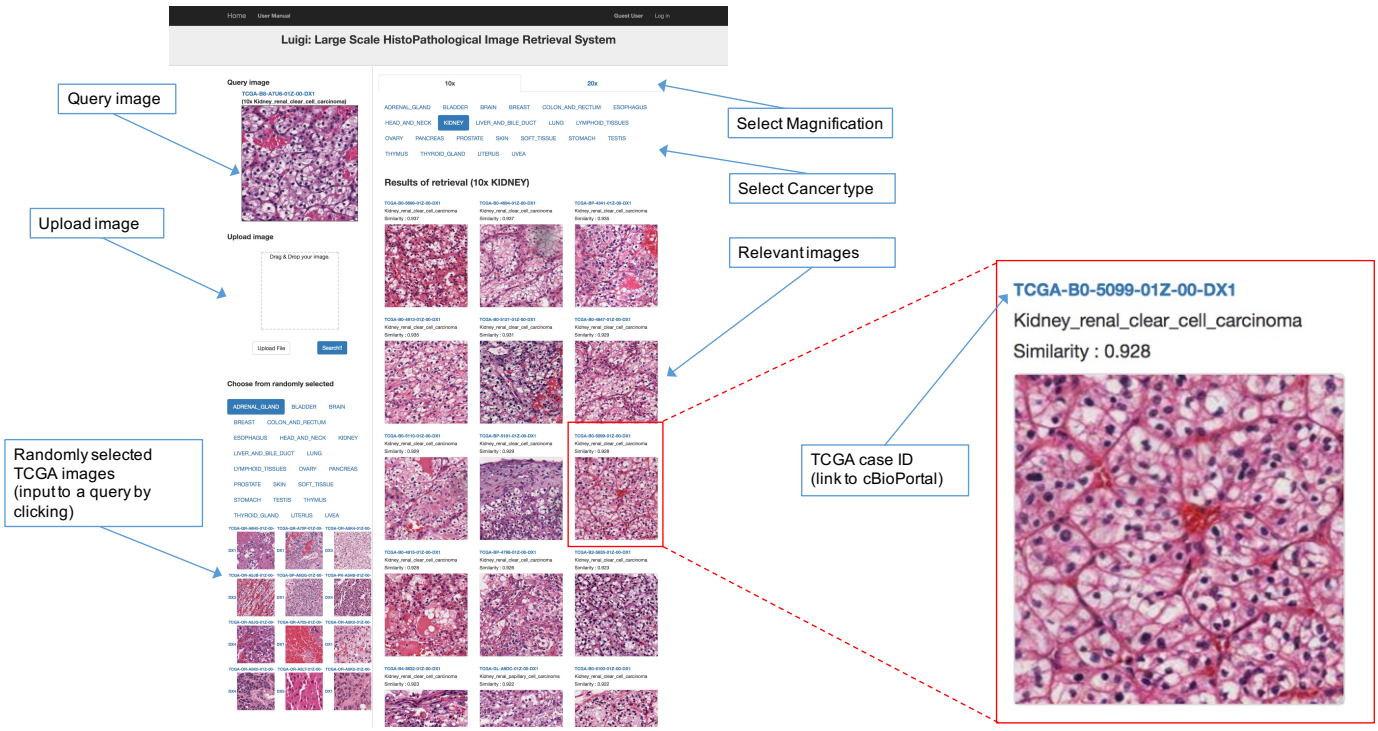
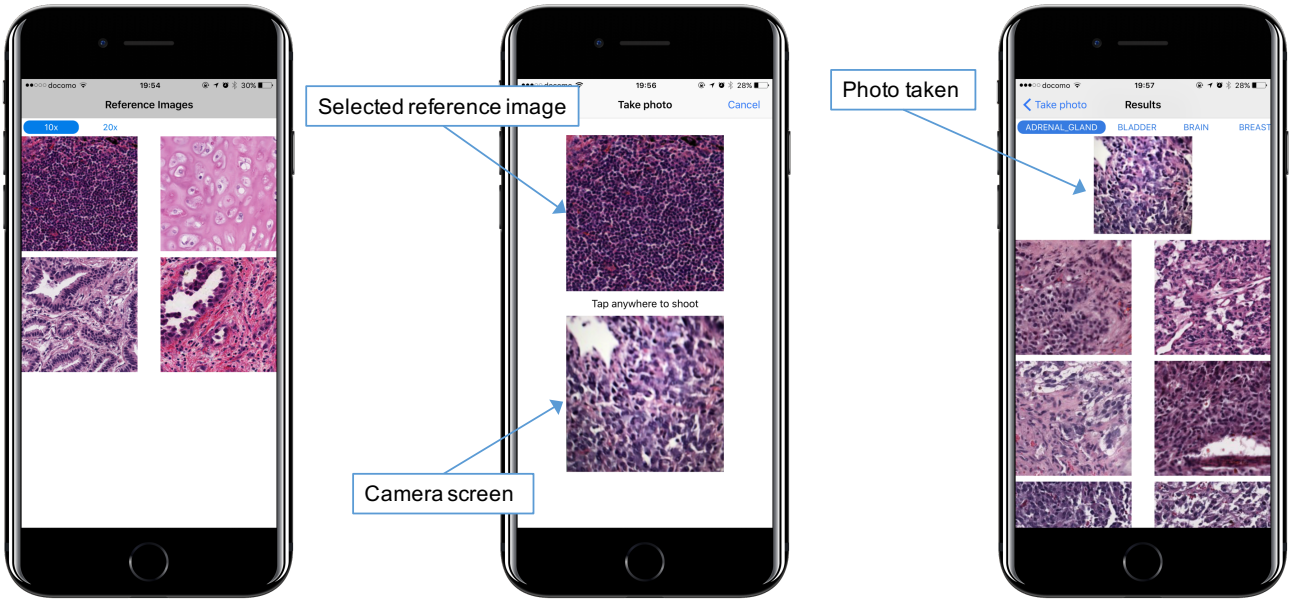


Figure 2. Graphical interface of the Luigi webserver.



1. Select a reference image and magnification

2. Take a photo of histopathological image

3. Display similar histopathological images by cancer type

Figure 3. Graphical interface and usage of the Luigi smartphone application.

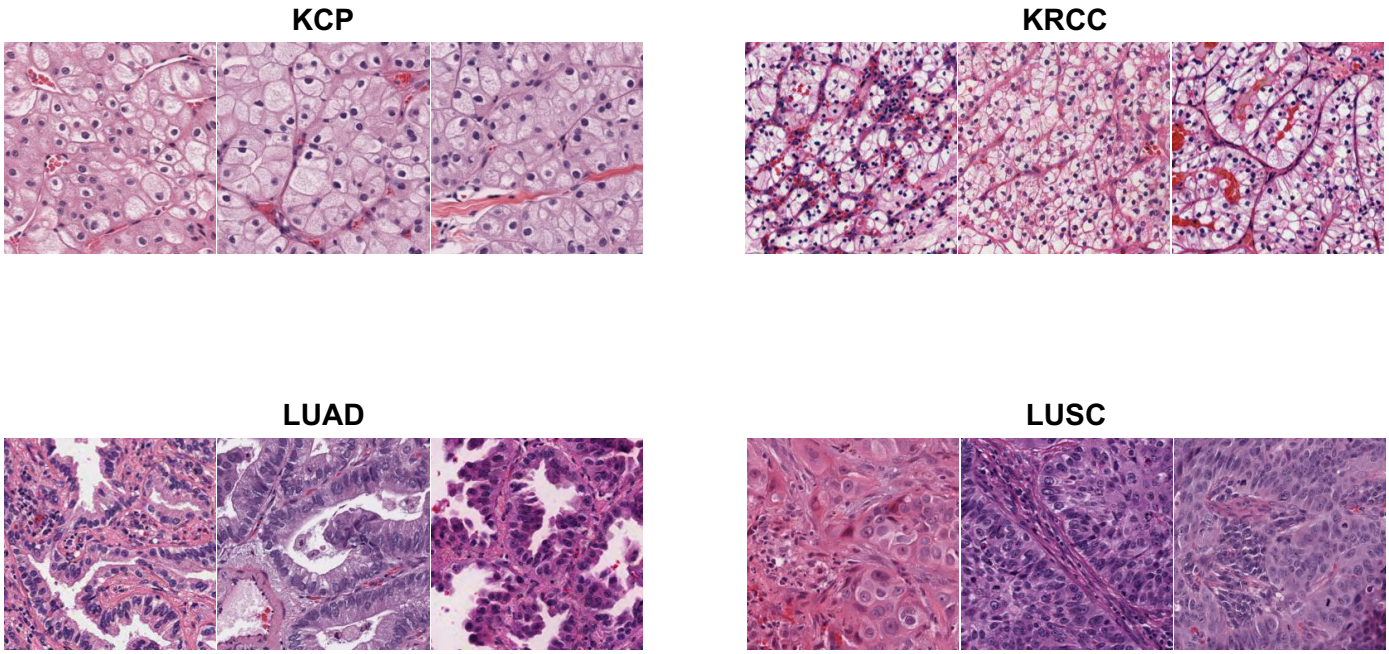


Figure 4. Query images of four cancer types used in the experiment.

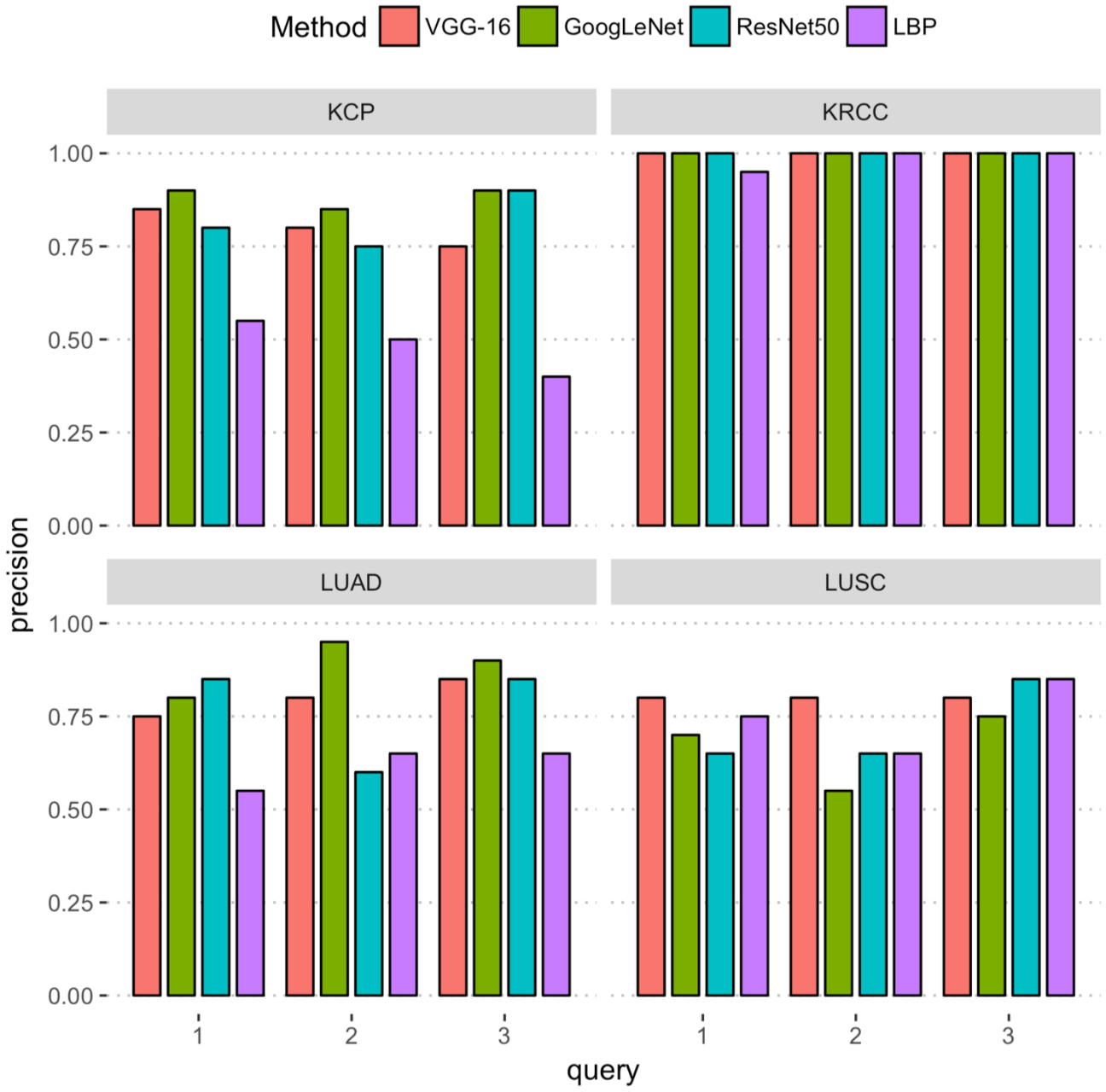


Figure 5. Retrieval results for a representative histopathological image in kidney cancer. A) kidney chromophobe renal cell carcinoma, B) kidney clear renal cell carcinoma. These results were obtained using the deep textures based on the VGG-16 model. The query image is shown on the top and the top five (six in kidney chromophobe renal cell carcinoma to show the incorrectly retrieved image) retrieved images are shown on the bottom. The case ID, cancer type case and cosine similarity to the query image of each case are shown above the retrieved image. Retrieved images of different cancer types matching the query image are highlighted in green boxes.

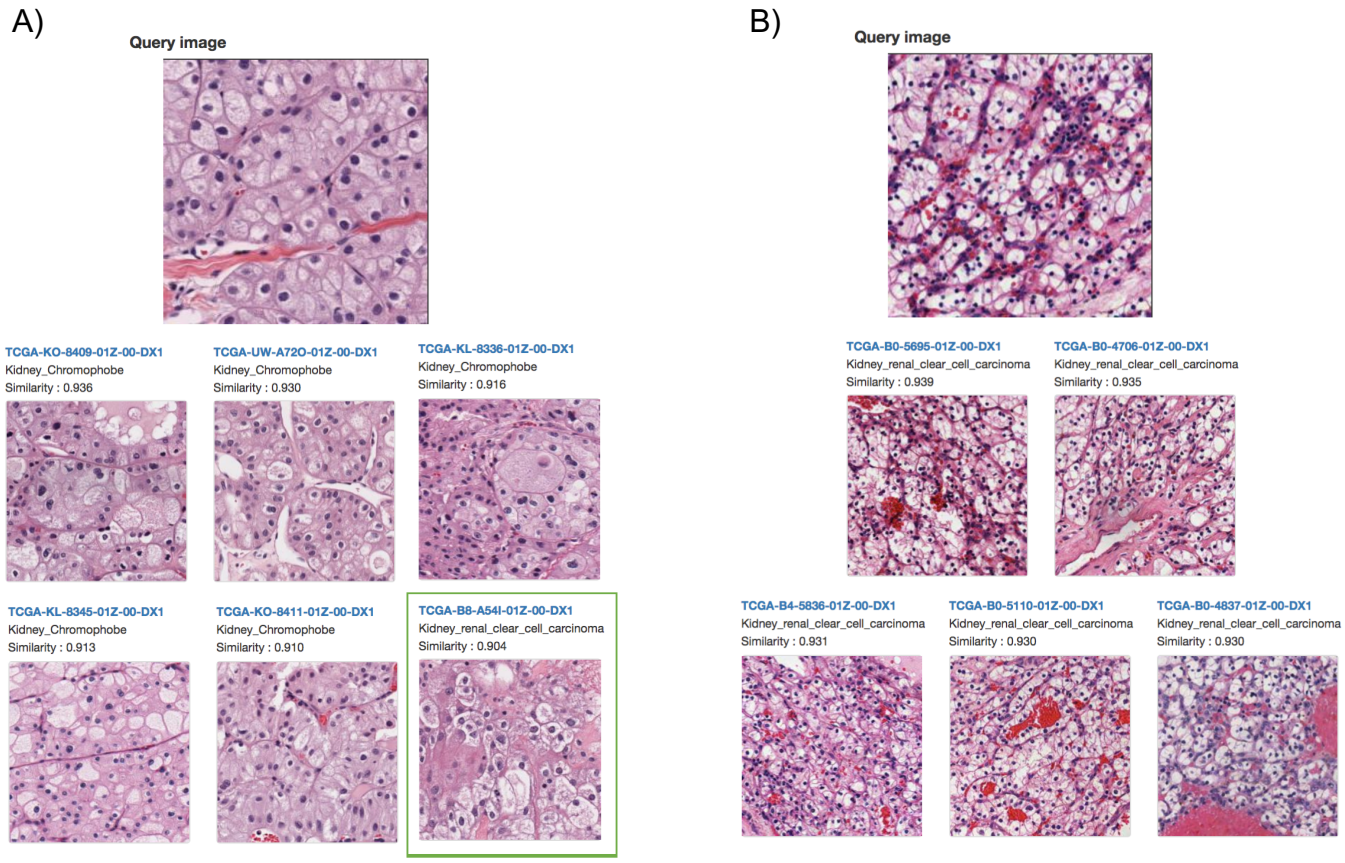


Figure 6. Retrieval results for a representative histopathological image in lung cancer. A) Lung adenocarcinoma, and B) lung squamous cell carcinoma. These results were obtained using the deep textures based on the VGG-16 model. The query image is shown on the top and the top five retrieved images are shown in the bottom. The case ID, cancer type case and cosine similarity to query image of each case are shown above the retrieved image. Retrieved images of different cancer types matching the query image are highlighted in green boxes.

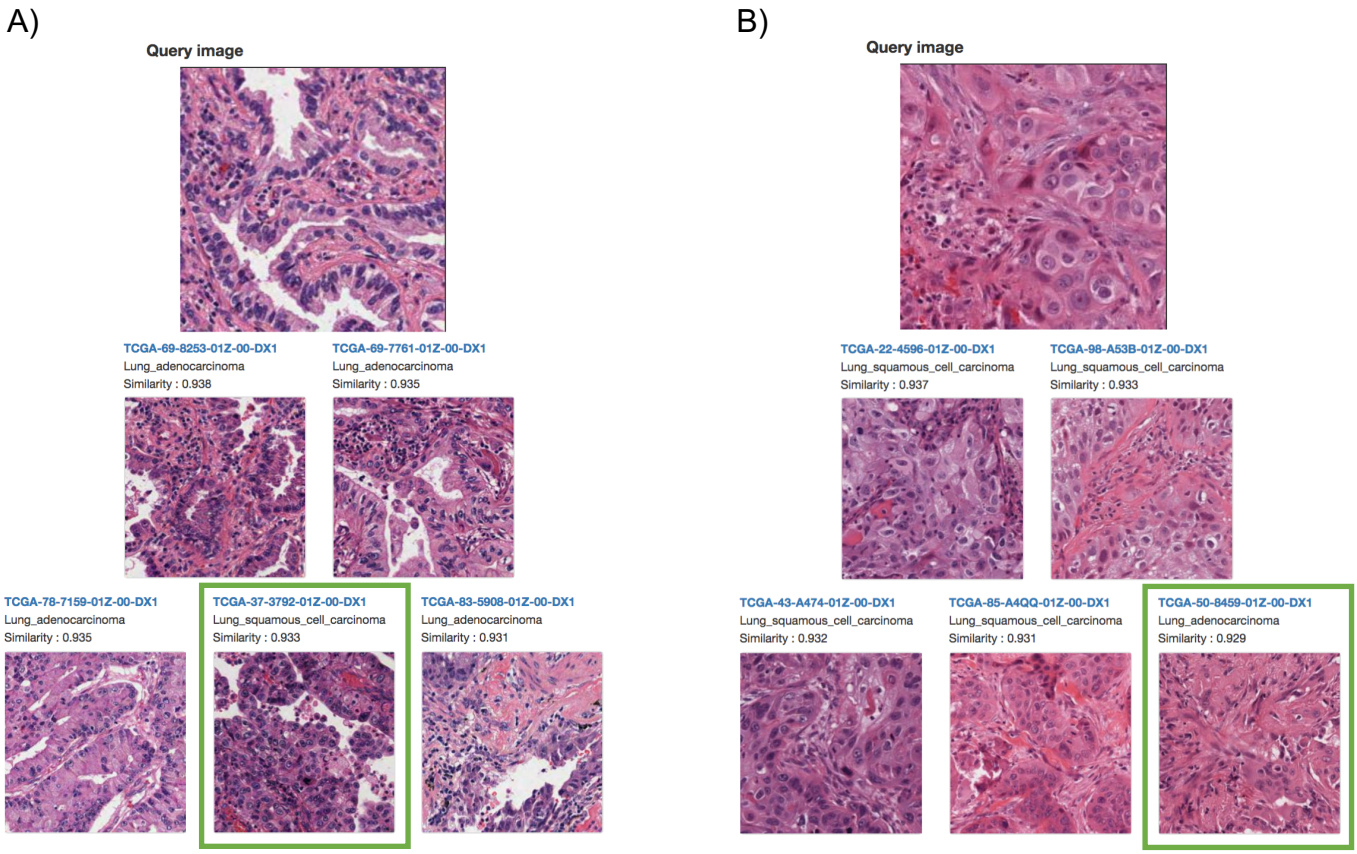


Figure 7. The precision of Luigi and LBP on The Cancer Genome Atlas dataset. Precision of Luigi based on VGG-16, GoogLeNet and ResNet50, and LBP are compared.

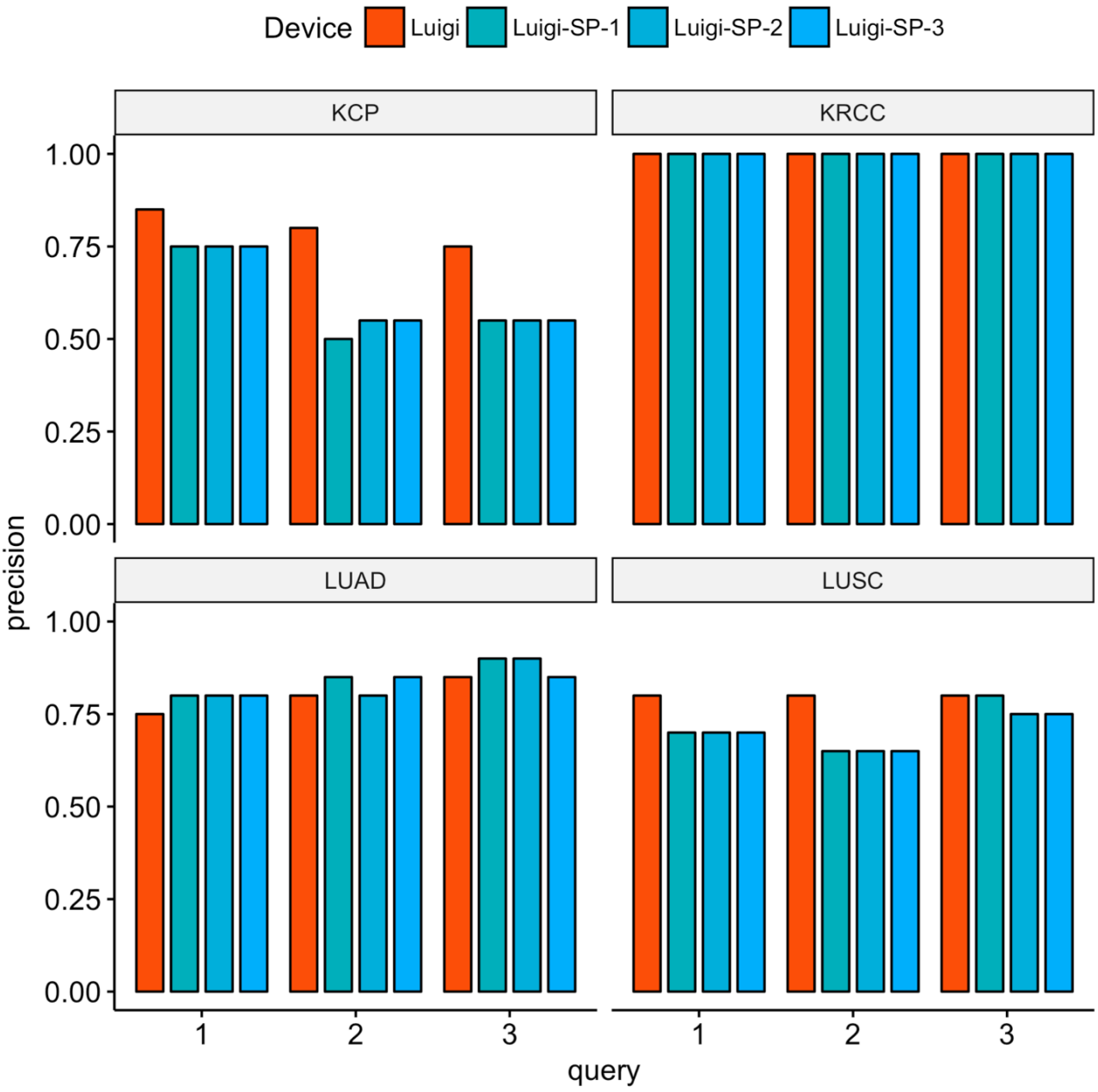


Figure 8 The precision of original image as the query and the images captured with smartphone camera on The Cancer Genome Atlas dataset. Precision of the original images as the query and images captured three times with the Luigi smartphone application (Luigi-SP) are compared. The VGG-16-based model was used in this experiment.

Table 1.

Cancer type	Primary site	No. of WSIs	Total number of patches		
			10x	20x	all
Breast invasive carcinoma	breast	877	389063	399085	788148
Glioblastoma multiforme	brain	843	359962	297098	657060
Brain Lower Grade Glioma	brain	839	362770	394721	757491
Lung adenocarcinoma	lung	531	236504	250193	486697
Lung squamous cell carcinoma	lung	513	241379	247785	489164
Sarcoma	soft tissue	459	223421	225394	448815
Uterine Corpus Endometrial Carcinoma	uterus	436	212300	214589	426889
Thyroid carcinoma	thyroid gland	386	185600	189073	374673
Head and neck squamous cell carcinoma	head and neck	359	170667	175926	346593
Skin cutaneous melanoma	skin	350	162835	166252	329087
Bladder urothelial carcinoma	bladder	346	167386	170306	337692
Colon adenocarcinoma	colon and rectum	344	163231	159588	322819
Prostate adenocarcinoma	prostate	334	150099	161509	311608
Kidney renal clear cell carcinoma	kidney	302	130550	180998	311548
Stomach adenocarcinoma	stomach	301	136517	136087	272604

Liver hepatocellular carcinoma	liver and bile duct	295	143196	146434	289630
Kidney renal papillary cell carcinoma	kidney	233	106300	112606	218906
Cervical squamous cell carcinoma and endocervical adenocarcinoma	uterus	215	91273	102470	193743
Testicular germ cell tumors	testis	206	99143	101736	200879
Adrenocortical carcinoma	adrenal gland	167	83040	83257	166297
Pancreatic adenocarcinoma	pancreas	158	71518	76836	148354
Pheochromocytoma and paraganglioma	adrenal gland	147	71778	72902	144680
Thymoma	thymus	131	61977	63728	125705
Rectum adenocarcinoma	colon and rectum	125	59109	55031	114140
Esophageal carcinoma	esophagus	114	53037	55985	109022
Kidney chromophobe	kidney	93	43960	46075	90035
Ovarian serous cystadenocarcinoma	ovary	84	40961	42002	82963
Uterine carcinosarcoma	uterus	73	35274	35391	70665
Mesothelioma	lung	71	30656	30861	61517
Uveal melanoma	uvea	55	27164	24437	51601
Lymphoid neoplasm diffuse large B-cell lymphoma	lymphoid tissues	33	14875	14852	29727

Cholangiocarcinoma	liver and bile duct	30	14896	15000	29896

5