

Image Analysis with R: a Review

Stefan Rödiger, Hinrich Winther and Michał Burdukiewicz

Abstract Management, display, and processing of biological and medical imaging data is an important task in life sciences and medical research. R is a powerful cross-platform, which can handle most task statistical computing tasks in the same environment. The aim of this mini-review is to give a brief overview about image processing software for the R statistical computing environment. When it comes to image analysis, R may appear to provide only few tools on the first sight. However, a systematic analysis of the existing packages shows that a huge potential for numerous applications.

Introduction

- Digital image processing?
- Where is it used?

There are numerous software tools that have been made available for digital image acquisition and processing (Wiesmann et al., 2015; Chieco et al., 2013). The scientific background of the users is deliberately broad. This includes biologists, biostatisticians, physicians and others, who fundamentally make use of the similar image analysis techniques. From discussion with peers we learned that knowledge of image processing is gained by self-study. In particular, learning of several programming languages may hamper the scientist to focus on their scientific aim. R (R Core Team, 2012) is *de facto* the *lingua franca* of statistical bioinformatics and therefore used in numerous research disciplines (Rödiger et al., 2015a). It is a powerful tool for statistical data analysis. It comes to no surprise that software packages for digital image processing have been implemented (Frery and Perciano, 2013). In this review, we give a current overview of the R ecosystem about which software packages exist and which deficits they may expose in comparison to other software packages. We aimed to aggregate information about R packages available on CRAN, Bioconductor (Gentleman et al., 2004), RForge or github.

Analysis of 2D and 3D digital images is a bridge technology which has been used to unravel quantitative and qualitative gene expression data (mRNA and protein level), cellular interactions and diagnostic data. In particular, immunofluorescence images were analyzed in relation to cell structures, tissues and organs (Chieco et al., 2013; Rödiger et al., 2013; Schierack et al., 2014; Willitzki et al., 2012).

We performed two image processing case studies where we applied selected packages for immunofluorescence image analysis and RMI data.

There are numerous software packages for the analysis of image data (Wiesmann et al., 2015). However, R is quite functional when it comes to digital image analysis.

Image processing capabilities of Cell-ID and data analysis by the statistical programming framework R for quantifying various cellular features (e.g., volume, total and subcellular fluorescence localization) from sets of microscope images of individual cells (Bush et al., 2012)

Give me a title

General image processing and analysis

(Tabelow et al., 2012, 2014)

Murrell (Murrell, 2011) **mmand** (Clayden, 2016)

CRAN provides well established packages. These are **jpeg**, **png** and **tiff** to read, write and display bitmap JPEG, PNG and TIFF images. The development of the **ripa** (Perciano and Frery, 2014) package was started in 2005 by Talita Perciano. This package can be used to process and analyse RGB, LAN (multispectral) and AVIRIS (hyperspectral) images. Recent advances of **ripa** make it a promising tool for analysis of large datasets. The vast amount of image data is becoming more and more an essential part of Big Data analysis pipelines. R is among the frequently used for data mining and analysis. It comes to no surprise the commercial and non-commercial entities make heavy use of R (Chen et al., 2014). **EBImage** (Pau et al., 2010) is presumably the most comprehensive package and the foundation for many other R packages in the context of microscopy-based cellular assays (Gowen, 2015). This package offers tools to transform (e.g., rotate) the images, segment object (e.g., cells) and extract quantitative descriptors. The early version of **EBImage** used the Magick++ interface to the ImageMagick image processing library (Sklyar and Huber, 2006).

General image processing and analysis

This section may contain a figure such as Figure 1.

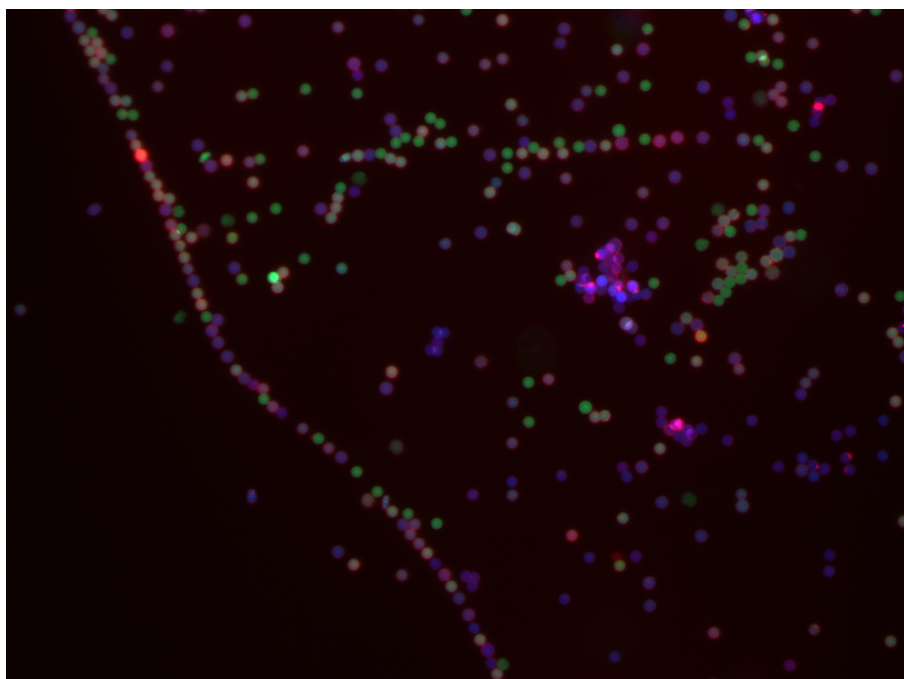


Figure 1: The logo of R.

segmentation

(Holmes et al., 2009)

adimpro is a package for manipulation of digital images and the Propagation Separation approach for smoothing digital images (Polzehl and Tabelow, 2007a). For example, image analysis is used for the detection and quantification of cell patterns and array technologies like microarrays and bead-based assays (Rödiger et al., 2013; Willitzki et al., 2012, 2013; Dunning et al., 2006). Several software packages have been developed. *imageJ* belongs to the most used and cited tools. When it comes to R numerous packages exist, which can be readily integrated in the analysis routines (Frery and Perciano, 2013).

The accuracy of image segmentation is a critical step in a computer-aided diagnosis systems. The recognition of mitotic cells and the classification of fluorescent patterns is heavily dependent on this. Immunofluorescent images of cell, such as Hep-2, exhibit a high variability due a wide range of staining patterns and intensity levels (FIGURES OF CELLS), the presence of mitotic cells and artifacts. The later may be caused by uneven illumination and photo-bleaching effects (Tonti et al., 2015).

Intensity inhomogeneity (bias field) is a common artefact in magnetic resonance (MR) images, which hinders successful automatic segmentation. (Ivanovska et al., 2016)

```
x <- 1:10
result <- myFunction(x)
```

Applications

Examples

Statistical analysis of functional magnetic resonance imaging data

Statistical analysis of functional magnetic resonance imaging (fMRI) data is a non-invasive neuroimaging technique.

largely used in clinical routine and advanced brain research

AnalyzefMRI (Marchini, 2002; Bordier et al., 2011) and **fmri** (Polzehl and Tabelow, 2007b) and are packages for the analysis of Magnetic Resonance Imaging (MRI) and functional Magnetic Resonance Imaging (fMRI) data, respectively.

Eventually these de

Others include **dcmriS4** (Dunning et al., 2006; Frery and Perciano, 2013).

Analysis of fluorescence image data

CRImage package (Failmezger et al., 2012) for tumor image analysis

Analysis of microbead assays

(Rödiger et al., 2014; Rödiger et al., 2013)

```
# Detailed installation instructions are available at
# http://bioconductor.org/packages/EBImage/

# Install the EBImage from Bioconductor
source("http://bioconductor.org/biocLite.R")
biocLite("EBImage")

# Load the EBImage package
require(EBImage)
```

Graphical User Interface for Digital Image Analysis

There exist also R graphical user interfaces (Rödiger et al., 2012) which can be used for image processing. Bio7 is an integrated development environment based on the Eclipse Rich Client Platform (RCP). The main purpose of this tool is the modeling and analysis of ecological systems. However, Bio7 is not restricted to this discipline. The application contains GUIs and plugins for simulation and analysis tasks. Interestingly, one of these plugins is an adaption image application ImageJ and another is available for a bidirectional Java connection to R. This means that data can be transferred from and to ImageJ and R.

Performance

Michał, would you like to take this section?

Requirements for recent research include the rapid processing of massive amounts of image data (Mega to Tera byte scale) that modern technologies (e.g., microscopes, MRI scanner) produce nowadays. Preferably, affordable personal desktop computers should be usable. R has several disadvantages when it comes to memory management and GPU and CPU usage ...

Summary

Many scientist are using R. However, it appears that only few make use of the image analysis tools currently available. We would like to raise awareness for the fact that R provides sophisticated packages for digital image analysis. The central advantage is that all analysis is conducted within the same environment accross most platforms. Added values for the user are that there is less need to learn a new programming languages and that all analysis can be performed in a consistent and cross-platform environment. Table 1 gives an overview of R packages currently available. We found that functions from the packages can be easily combined to conduct manipulations and analysis (object transformations, measurements, object counting, grey level statisitcs, binarization, etc) at advanced levels. The organization of such functions in coustumized packages is straigthforward even for more complex analysis functions.

(Rödiger et al., 2015b) or combined assays of microbeads and cells (Scholz et al., 2015; Grossmann et al., 2016)

Package	Main function	Comment	Source
adimpro	×	×	×
AnalyzeFMRI	×	×	×
CRImage	×	×	Bioconductor
dcemriS4	×	×	CRAN
EBImage	fancy stuff	well maintained	Bioconductor
jpeg	×	×	CRAN
png	×	×	CRAN
ripa	×	×	×
tiff	×	×	CRAN
videoTools	video and image analysis	-	RForge

Table 1: R packages. *videoTools*: <http://www.rforge.net/videoTools/files/>.

While evaluation the packages we ended up with experiences regarding the installation and maintenance. Basically all installation were well instructed. For example, the installation of sophisticated packages like **EBImage** is well documented. However, many packages depend on a number of third-party packages. This applies also to **EBImage** which requires tiff and fftwtools.

Others and we recommend to make use of packages for reproducible research (Rödiger et al., 2015a).

There exist several GUI technologies for R which make it possible to integrate R code into an easy to master point-and-click interface.

Bibliography

- C. Bordier, M. Dojat, and P. L. De Micheaux. Temporal and spatial independent component analysis for fMRI data sets embedded in the AnalyzeFMRI R package. *Journal of Statistical Software*, 44(9): 1–24, 2011. URL <https://hal.archives-ouvertes.fr/inserm-00659425/>. [p3]
- A. Bush, A. Chernomoretz, R. Yu, A. Gordon, and A. Colman-Lerner. Using Cell-ID 1.4 with R for Microscope-Based Cytometry. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, CHAPTER:Unit14.18, Oct. 2012. ISSN 1934-3639. doi: 10.1002/0471142727.mb1418s100. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3485637/>. [p1]
- M. Chen, S. Mao, and Y. Liu. Big Data: A Survey. *Mobile Networks and Applications*, 19(2):171–209, Jan. 2014. ISSN 1383-469X, 1572-8153. doi: 10.1007/s11036-013-0489-0. [p1]
- P. Chieco, A. Jonker, B. A. De Boer, J. M. Ruijter, and C. J. F. Van Noorden. Image Cytometry: Protocols for 2d and 3d Quantification in Microscopic Images. *Progress in Histochemistry and Cytochemistry*, 47(4):211–333, Jan. 2013. ISSN 0079-6336. doi: 10.1016/j.proghi.2012.09.001. URL <http://www.sciencedirect.com/science/article/pii/S007963361200037X>. [p1]
- J. Clayden. *mmmand: Mathematical Morphology in Any Number of Dimensions*. 2016. URL <https://CRAN.R-project.org/package=mmmand>. R package version 1.3.0. [p1]
- M. Dunning, M. Smith, N. Thorne, and S. Tavaré. beadarray: An R Package to Analyse Illumina BeadArrays. *R News*, 6(5):17–23, Dec. 2006. URL <http://CRAN.R-project.org/doc/Rnews/>. [p2, 3]
- H. Failmezger, Y. Yuan, O. Rueda, and F. Markowetz. *CRImage: CRImage a package to classify cells and calculate tumour cellularity*. 2012. R package version 1.18.0. [p3]
- A. C. Frery and T. Perciano. *Introduction to Image Processing Using R*. SpringerBriefs in Computer Science. Springer London, London, 2013. ISBN 978-1-4471-4949-1 978-1-4471-4950-7. doi: 10.1007/978-1-4471-4950-7. [p1, 2, 3]
- R. C. Gentleman, V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. Yang, and J. Zhang. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10):R80, Sept. 2004. ISSN 1465-6906. doi: 10.1186/gb-2004-5-10-r80. URL <http://genomebiology.com/2004/5/10/R80/abstract>. 07026. [p1]

- A. A. Gowen. Near infrared hyperspectral image analysis using R. Part 6: image processing using EBIImage. *NIR news*, 26(3):20–21, 2015. ISSN 0960-3360. doi: 10.1255/nirn.1526. [p1]
- K. Grossmann, N. Röber, R. Hiemann, S. Rödiger, P. Schierack, D. Reinhold, M. W. Laass, K. Conrad, and D. Roggenbuck. Simultaneous detection of celiac disease-specific IgA antibodies and total IgA. *Autoimmunity Highlights*, 7(1):1–10, Jan. 2016. ISSN 2038-0305, 2038-3274. doi: 10.1007/s13317-016-0073-2. URL <http://link.springer.com/article/10.1007/s13317-016-0073-2>. [p]
- S. Holmes, A. Kapelner, and P. P. Lee. An Interactive Java Statistical Image Segmentation System: GemIdent. *Journal of statistical software*, 30(10), June 2009. ISSN 1548-7660. URL <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3100170/>. [p2]
- T. Ivanovska, R. Laqua, L. Wang, A. Schenk, J. H. Yoon, K. Hegenscheid, H. Völzke, and V. Liebscher. An efficient level set method for simultaneous intensity inhomogeneity correction and segmentation of MR images. *Computerized Medical Imaging and Graphics*, 48:9–20, Mar. 2016. ISSN 0895-6111. doi: 10.1016/j.compmedimag.2015.11.005. URL <http://www.sciencedirect.com/science/article/pii/S0895611115001779>. [p2]
- J. Marchini. AnalyzeFMRI: An R package for the exploration and analysis of MRI and fMRI datasets. *R News*, 2(1):17–23, Mar. 2002. URL <http://www.r-project.org/doc/2002-03/NEWS.html>. [p3]
- P. Murrell. Raster Images in R Graphics. *The R Journal*, 3(1):48–54, June 2011. URL http://journal.r-project.org/archive/2011-1/RJournal_2011-1_Murrell.pdf. [p1]
- G. Pau, F. Fuchs, O. Sklyar, M. Boutros, and W. Huber. EBIImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics*, 26(7):979–981, 2010. doi: 10.1093/bioinformatics/btq046. [p1]
- T. Perciano and A. C. Frery. ripa: R Image Processing and Analysis, May 2014. URL <https://cran.r-project.org/web/packages/ripa/index.html>. [p1]
- J. Polzehl and K. Tabelow. Adaptive Smoothing of Digital Images: The R Package adimpro. *Journal of Statistical Software*, 19(1):1–17, 2007a. ISSN 1548-7660. URL <http://www.jstatsoft.org/index.php/jss/article/view/v019i01>. [p2]
- J. Polzehl and K. Tabelow. fmri: A Package for Analyzing fmri Data. *R News*, 7(2):13–17, Oct. 2007b. URL <http://www.r-project.org/doc/2007-10/NEWS.html>. [p3]
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL <http://www.R-project.org/>. ISBN 3-900051-07-0. [p1]
- S. Rödiger, P. Schierack, A. Böhm, J. Nitschke, I. Berger, U. Frömmel, C. Schmidt, M. Ruhland, I. Schimke, D. Roggenbuck, W. Lehmann, and C. Schröder. A highly versatile microscope imaging technology platform for the multiplex real-time detection of biomolecules and autoimmune antibodies. *Advances in Biochemical Engineering/Biotechnology*, 133:35–74, 2013. ISSN 0724-6145. doi: 10.1007/10_2011_132. [p1, 2, 3]
- S. Rödiger, T. Friedrichsmeier, P. Kapat, and M. Michalke. RKWard: A Comprehensive Graphical User Interface and Integrated Development Environment for Statistical Analysis with R. *Journal of Statistical Software*, 49(9):1–34, 2012. ISSN 1548-7660. URL <http://www.jstatsoft.org/v49/i09>. [p3]
- S. Rödiger, C. Liebsch, C. Schmidt, W. Lehmann, U. Resch-Genger, U. Schedler, and P. Schierack. Nucleic acid detection based on the use of microbeads: a review. *Microchimica Acta*, 181(11-12): 1151–1168, Aug. 2014. ISSN 0026-3672, 1436-5073. doi: 10.1007/s00604-014-1243-4. URL <http://link.springer.com/article/10.1007/s00604-014-1243-4>. [p]
- S. Rödiger, M. Burdukiewicz, K. A. Blagodatskikh, and P. Schierack. R as an Environment for the Reproducible Analysis of DNA Amplification Experiments. *The R Journal*, 7(2):127–150, 2015a. URL <http://journal.r-project.org/archive/2015-1/RJ-2015-1.pdf>. 00000. [p1, 4]
- S. Rödiger, T. Kramer, U. Frömmel, J. Weinreich, D. Roggenbuck, S. Guenther, K. Schaufler, C. Schröder, and P. Schierack. Intestinal *Escherichia coli* colonization in a mallard duck population over four consecutive winter seasons. *Environmental Microbiology*, pages 1–10, 2015b. ISSN 1462-2920. doi: 10.1111/1462-2920.12807. URL <http://onlinelibrary.wiley.com/doi/10.1111/1462-2920.12807/abstract>. 00000. [p]

- P. Schierack, S. Rödiger, R. Kolenda, R. Hiemann, E. Berger, K. Grzymajło, A. Swidsinski, T. Juretzek, D. Meissner, K. Mydlak, D. Reinhold, L. K. Nolan, and D. Roggenbuck. Species-specific and pathotype-specific binding of bacteria to zymogen granule membrane glycoprotein 2 (GP2). *Gut*, pages gutjnl-2014-307854, July 2014. ISSN , 1468-3288. doi: 10.1136/gutjnl-2014-307854. URL <http://gut.bmj.com/content/early/2014/07/29/gutjnl-2014-307854.00003>. [p1]
- J. Scholz, K. Grossmann, I. Knütter, R. Hiemann, M. Sowa, N. Röber, S. Rödiger, P. Schierack, D. Reinhold, D. P. Bogdanos, P. L. Meroni, A. Radice, K. Conrad, and D. Roggenbuck. Second generation analysis of antinuclear antibody (ANA) by combination of screening and confirmatory testing. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 0(0):1–12, May 2015. ISSN 1437-4331. doi: 10.1515/cclm-2015-0083. URL <http://www.degruyter.com/view/j/cclm.ahead-of-print/cclm-2015-0083/cclm-2015-0083.xml>. [p]
- O. Sklyar and W. Huber. Image Analysis for Microscopy Screens. *R News*, 6(5):12–16, Dec. 2006. URL <http://CRAN.R-project.org/doc/Rnews/>. [p1]
- K. Tabelow, H. U. Voss, and J. Polzehl. Modeling the orientation distribution function by mixtures of angular central Gaussian distributions. *Journal of Neuroscience Methods*, 203(1):200–211, Jan. 2012. ISSN 0165-0270. doi: 10.1016/j.jneumeth.2011.09.001. URL <http://www.sciencedirect.com/science/article/pii/S0165027011005231>. [p1]
- K. Tabelow, J. Polzehl, and F. Anker. dti: Analysis of diffusion weighted imaging (DWI) data, Dec. 2014. URL <https://cran.r-project.org/web/packages/dti/index.html>. [p1]
- S. Tonti, S. D. Cataldo, A. Bottino, and E. Ficarra. An automated approach to the segmentation of HEp-2 cells for the indirect immunofluorescence ANA test. *Computerized Medical Imaging and Graphics*, 40:62–69, Mar. 2015. ISSN 0895-6111. doi: 10.1016/j.compmedimag.2014.12.005. URL <http://www.medicalimagingandgraphics.com/article/S089561111400202X/abstract>. [p2]
- V. Wiesmann, D. Franz, C. Held, C. Münzenmayer, R. Palmisano, and T. Wittenberg. Review of free software tools for image analysis of fluorescence cell micrographs. *Journal of Microscopy*, 257(1): 39–53, Jan. 2015. ISSN 1365-2818. doi: 10.1111/jmi.12184. URL <http://onlinelibrary.wiley.com/doi/10.1111/jmi.12184/abstract>. [p1]
- A. Willitzki, R. Hiemann, V. Peters, U. Sack, P. Schierack, S. Rödiger, U. Anderer, K. Conrad, D. P. Bogdanos, D. Reinhold, and D. Roggenbuck. New platform technology for comprehensive serological diagnostics of autoimmune diseases. *Clinical & developmental immunology*, 2012:284740, 2012. ISSN 1740-2530. doi: 10.1155/2012/284740. [p1, 2]
- A. Willitzki, S. Lorenz, R. Hiemann, K. Guttek, A. Goihl, R. Hartig, K. Conrad, E. Feist, U. Sack, P. Schierack, L. Heiserich, C. Eberle, V. Peters, D. Roggenbuck, and D. Reinhold. Fully automated analysis of chemically induced γ H2AX foci in human peripheral blood mononuclear cells by indirect immunofluorescence. *Cytometry. Part A: the journal of the International Society for Analytical Cytology*, 83(11):1017–1026, Nov. 2013. ISSN 1552-4930. doi: 10.1002/cyto.a.22350. [p2]

Stefan Rödiger (corresponding author)
orcid.org/0000-0002-1441-6512
Faculty of Natural Sciences
Brandenburg University of Technology Cottbus–Senftenberg
Senftenberg
Germany
Stefan.Roediger@b-tu.de

Hinrich Winther
Affiliation
Address
Country
author@work

Michał Burdukiewicz
University of Wrocław
Faculty of Biotechnology
Department of Genomics
Wrocław
Poland michalburdukiewicz@gmail.com