Estimating the width of uniform distribution under measurement errors

Mirta Benšić, Kristian Sabo, Safet Hamedović

M. Benšić, K. Sabo — Department of Mathematics, University of Osijek

S. Hamedović — Faculty of Metallurgy and Technology, University of Zenica

Firenze, Jun 2019

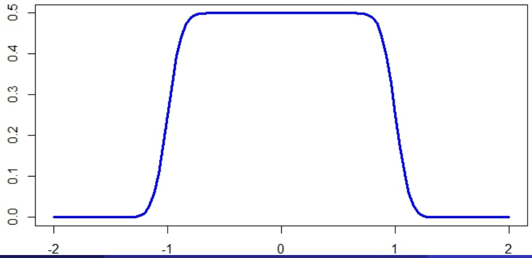
This work is supported by the Croatian Science Foundation through research grants IP-2016-06-6545

Convolution model

- $X = U + \varepsilon$
- U i ε independent, random
- U uniform, ε normal error

  Example: $U \sim \mathcal{U}(-a, a)$, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$

$$f_X(x; a, \sigma) = \frac{1}{2a}\left(F_\varepsilon\left(\frac{x+a}{\sigma}\right) - F_\varepsilon\left(\frac{x-a}{\sigma}\right)\right)$$



Applications in metrology

- Usually assumed that ε is normally distributed.
- Known as the Flatten–Gaussian distribution
- The basis for calculating the measurement uncertainty
  - Blázquez, J., García-Berrocal, A., Montalvo, C., Balbás, M. (2008). The coverage factor in a Flatten–Gaussian distribution, Metrologia 45, 503–506.
  - Fotowicz, P. (2014). Methods for calculating the coverage interval based on the Flatten–Gaussian distribution. Measurement, 55, 272-275

Applications in geology

- Cox, A.V., Dalrymple, G.B. (1967). Statistical analysis of geomagnetic reversal data and the precision of potassium-argon dating, J Geophys Res 72:2603–2614
- Agterberg, F.P. (1988). Quality of time scales – a statistical appraisal. In: Merriam, D.F. (ed) Current trends in geomathematics, Plenum, New York, pp 57–103
- Agterberg, F. P. (2014). Geomathematics: Theoretical foundations, applications and future developments. Quantitative geology and geostatistics (Vol. 18), Springer, Heidelberg.

- Ex. — for estimating the age of stage boundaries.

Applications in electrical engineering

- For modeling the transmission losses data
  Tolić, I., K. Miličević, N. Šuvak, I. Biondić (2017). Non-linear Least Squares and Maximum Likelihood Estimation of Probability Density Function of Cross-Border Transmission Losses, IEEE Transactions on Power Systems 33/2 (2018), 2230–2238

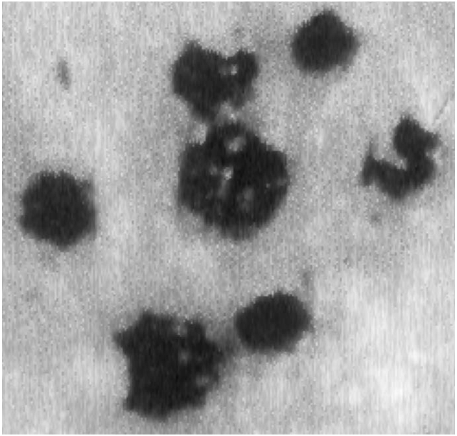
Fitting a line segment to noisy data

- Uniform latent variable and normal noise
  Davidov, O., Goldenshluger, A. (2004). Fitting a line segment to noisy data. Journal of Statistical Planing and Inference 119, 191-206.
- Linear structural relationship
  Chan, N.N. (1982). Linear structural relationships with unknown error variances, Biometrika, 69, No.1, 277–279

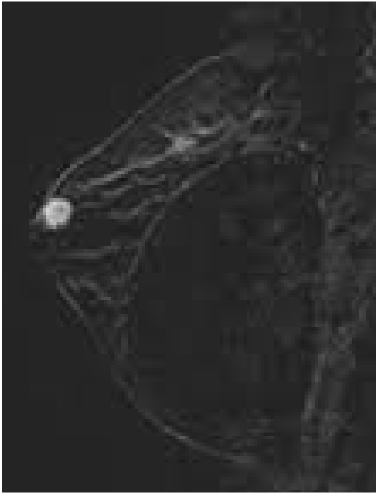
Estimating the size of an object in a noisy image

- M. B., K. Sabo, *Border estimation of a Two-dimensional Uniform Distribution if Data are Measured with Additive Error*, Statistics, 41 (2007), 4, 311–319.
- K. Sabo, M. B., *Border estimation of a disc observed with random errors solved in two steps*, Journal of Computational and Applied Mathematics, 229 (2009)

Black fungi colonies



Garlipp, T., Müller, C. H., Detection of linear and circular shapes in image analysis, Computational Statistics & Data Analysis 51 (2006), 1479–1490

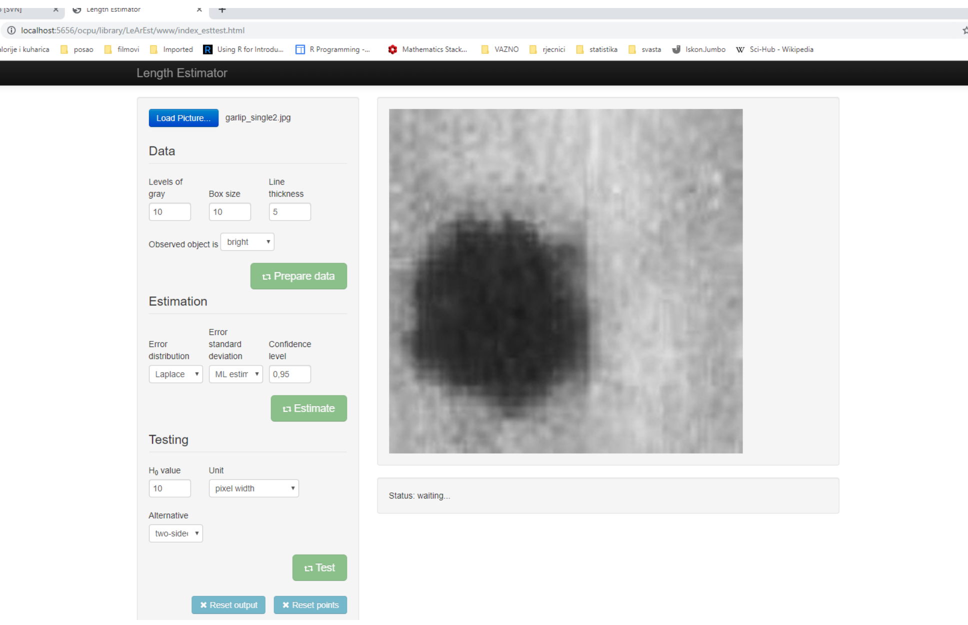


R package

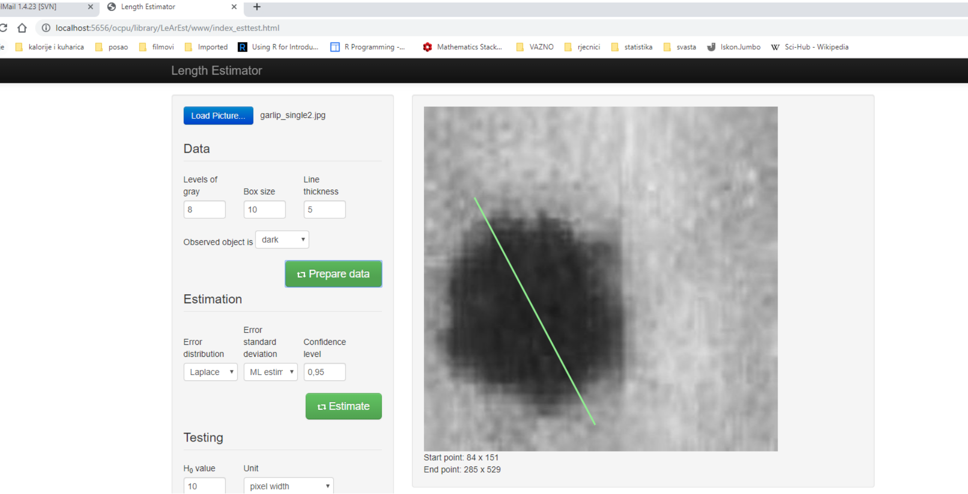
$https : //cran.r - project.org/web/packages/LeArEst/$

- M. Benšić, P. Taler, S. Hamedović, E.K. Nyarko, K. Sabo, *LeArEst: Length and Area Estimation from Data Measured with Additive Error*, The R Journal 9/2 (2017), 461-473
- Includes web application for estimating the size of an object from a noisy image
- Gaussian error model — tested in simulations

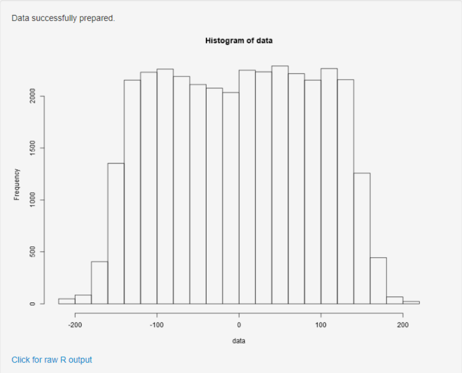
LeArEst


Length Estimator
localhost:5656/ocpu/library/LeArEst/www/index_esttest.html
Length Estimator
Load Picture...
garlip_single2.jpg
Data
Levels of gray
Box size
Line thickness
10
10
5
Observed object is
bright
Prepare data
Estimation
Error distribution
Error standard deviation
Confidence level
Laplace
ML estim
0,95
Estimate
Testing
H0 value
Unit
10
pixel width
Alternative
two-side
Test
Reset output
Reset points
Status: waiting...

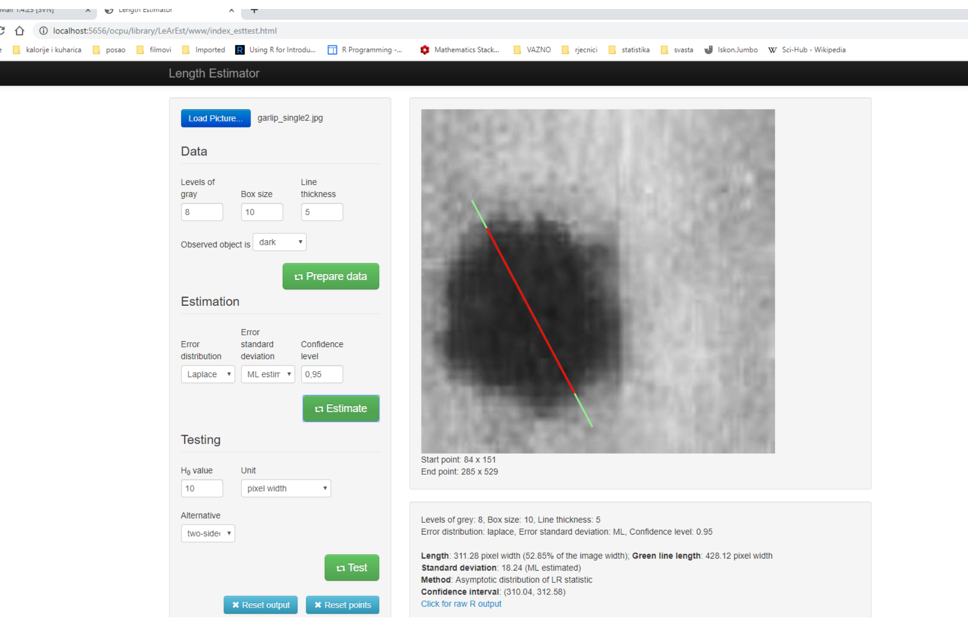
LeArEst


IMail 1.4.23 [SVN]
Length Estimator
localhost:5656/ocpu/library/LeArEst/www/index_esttest.html
kalorije i kuharica
posao
filmovi
Imported
Using R for Introdu...
R Programming -...
Mathematics Stack...
VAZNO
rjecnici
statistika
svasta
Iskon.Jumbo
Sci-Hub - Wikipedia
Length Estimator
Load Picture...
garlip_single2.jpg
Data
Levels of gray
8
Box size
10
Line thickness
5
Observed object is
dark
Prepare data
Estimation
Error distribution
Laplace
Error standard deviation
ML estim
Confidence level
0,95
Estimate
Testing
H0 value
10
Unit
pixel width
Start point: 84 x 151
End point: 285 x 529

LeArEst

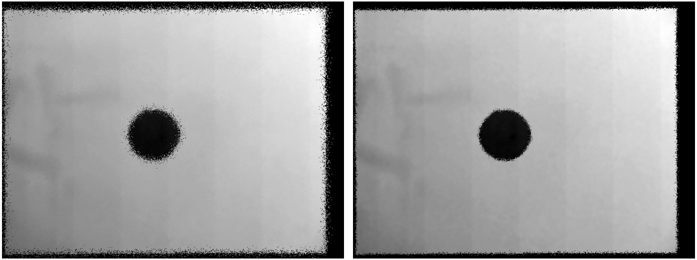

Data successfully prepared.
Histogram of data
Frequency
0
500
1000
1500
2000
-200
-100
0
100
200
data
Click for raw R output

LeArEst


localhost:5656/ocpu/library/LeArEst/www/index_esttest.html
Length Estimator
Load Picture...
garlip_single2.jpg
Data
Levels of gray
Box size
Line thickness
8
10
5
Observed object is
dark
Prepare data
Estimation
Error distribution
Error standard deviation
Confidence level
Laplace
ML estim
0,95
Estimate
Testing
H0 value
Unit
10
pixel width
Alternative
two-sided
Test
Reset output
Reset points
Start point: 84 x 151
End point: 285 x 529
Levels of grey: 8, Box size: 10, Line thickness: 5
Error distribution: laplace, Error standard deviation: ML, Confidence level: 0.95
Length: 311.28 pixel width (52.85% of the image width); Green line length: 428.12 pixel width
Standard deviation: 18.24 (ML estimated)
Method: Asymptotic distribution of LR statistic
Confidence interval: (310.04, 312.58)
Click for raw R output

Detection of the circle edge in a noisy image

Image of the can, $r = 51.5mm$, $area = 8332.3mm^2$



Detection of the circle edge in a noisy image

- Hough transformation
  P. V. Hough,*Method and means for recognizing complex patterns*. Patent U.S. Patent No. 3,069,654, 1962.
  R. D. Duda i P. E. Hart, *Use ofthe Hough transform to detectlinesand curvesin pictures*, Commun. ACM, 15, br. 1, 1972.
  A. Rosenfeld, *Picture processing by computer*, ACM Computing Surveys (CSUR), svez. 1, br. 3, pp. 147-176, 1969.
- EDCirlces
  C. Akinlar i C. Topal, *EDCircles: A real-time circle detector with a false detection control*, Pattern Recognition, 46, br. 3, pp. 725-740, 2013.
- Fornaciari
  M. Fornaciari, A. Prati i R. Cucchiara, *A fast and effective ellipse detector for embedded vision applications*, Pattern Recognition, 47, br. 11, pp. 3693-3708, 2017.

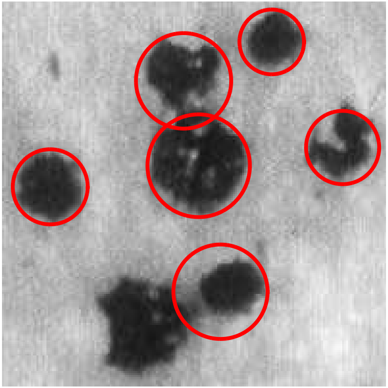
Simulation results — can example

- $r = 51.5mm$, $area = 8332.3mm^2$
- 20 simulations for each sd
- RMSE for area estimation

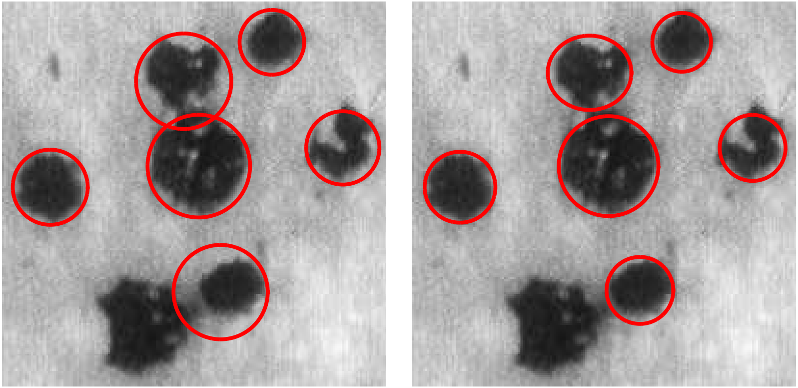
$sd/r(\%)$	LeArEst	Hough	EDCircles	Fornaciari
0	229	1858	681	
5	587	1353	723	
13	523	1365	852	
26	243	2315	659	
39	289	1955	755	

- For 1000 simulations we have similar results but withouth EDCircles.

Problem with real images



Laplace error model — improvement



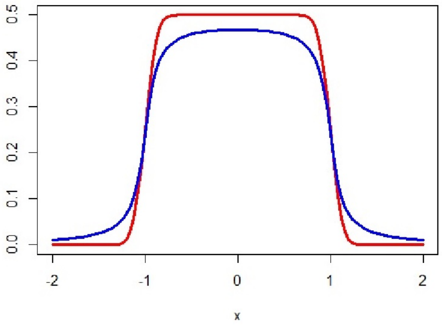
General one-dimensional model

- $X = U + \sigma\varepsilon$
- $f_U(t) = \frac{1}{2a} I_{[-a,a]}(t)$
- ε absolutely continuous random variable with distribution F_ε and density function f_ε, which is even.
- $$\Rightarrow \quad f_X(t) = \frac{1}{2a}[F_\varepsilon(\frac{a+t}{\sigma}) - F_\varepsilon(\frac{a-t}{\sigma})]$$
- $\varepsilon \sim \mathcal{N}(0,1)$

  $$f_X(t|a,\sigma) = \frac{1}{2a\sqrt{2\pi}} \int_{-\frac{a+t}{\sigma}}^{\frac{a-t}{\sigma}} e^{-\frac{u^2}{2}} du$$
- $\varepsilon \sim \mathcal{L}(\lambda)$, $\sigma = 1$

  $$f_X(t|a,\lambda) = \begin{cases} \frac{1}{2a} \sinh \frac{a}{\lambda} e^{-\frac{|t|}{\lambda}}, & a \in (0,|t|], \\ \frac{1}{2a}\left(1 - e^{-\frac{a}{\lambda}} \cosh \frac{t}{\lambda}\right), & a \in (|t|,\infty). \end{cases}$$

General one-dimensional model - density



red — normal error
blue — Student's $t(1)$ error

Maximum likelihood estimation

- Data $(x_1, \dots, x_n)$, log-likelihood:

$$l(a) = -n \log 2a + \log \sum_{i=1}^{n} \left[F_\varepsilon \left(\frac{x_i + a}{\sigma} \right) - F_\varepsilon \left(\frac{x_i - a}{\sigma} \right) \right]$$

- Differentiable function (model with absolutely continuous error).
- Optimization should be easy with a good initial approximation that is not difficult to recognize in applications.
- Regularity conditions???

MLE - regularity conditions

- Easy to check sufficient conditions
- S. Hamedović, MB, K. Sabo, *Estimating the width of a uniform distribution under symmetric measurement errors*, submitted 2019
- Examples: Normal, Logistic, Student's $t(\nu)$, $\nu \geq 1$
- Student's distribution with small degrees of freedom could be a particularly good choice for the error. (We can adjust the number of degrees of freedom to the amount of outliers in the data.)

General regular model

ML confidence intervals

Fisher information

$$I(a) = \frac{-1}{a^2} + \frac{1}{2a\sigma^2} \int_{-\infty}^{\infty} \frac{\left(f_\varepsilon\left(\frac{x+a}{\sigma}\right) + f_\varepsilon\left(\frac{x-a}{\sigma}\right)\right)^2}{F_\varepsilon\left(\frac{x+a}{\sigma}\right) - F_\varepsilon\left(\frac{x-a}{\sigma}\right)}\, dx$$

$$\left(\hat{a}_{ML} - \frac{z_{1-\alpha/2}}{\sqrt{nI(\hat{a}_{ML})}}, \hat{a}_{ML} + \frac{z_{1-\alpha/2}}{\sqrt{nI(\hat{a}_{ML})}}\right)$$

General model

LR test

$$H_0 : a = a_0$$

$$H_1 : a \neq a_0$$

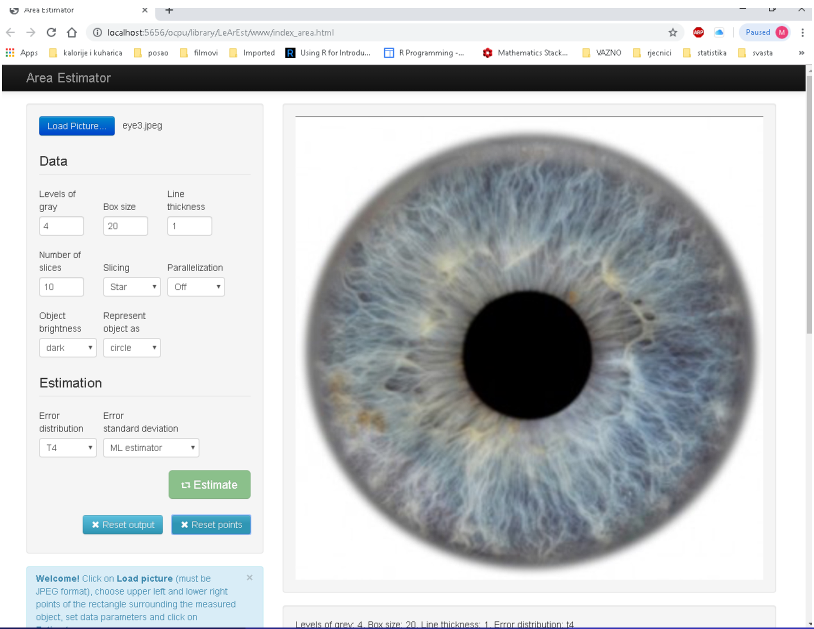
Critical region (significance level α, L likelihood)

$$\left\{ \mathbf{y} \middle| -2 \log \frac{L(a_0; \mathbf{y})}{L(\hat{a}_{ML}; \mathbf{y})} \geq \chi^2_1(1-\alpha) \right\}$$

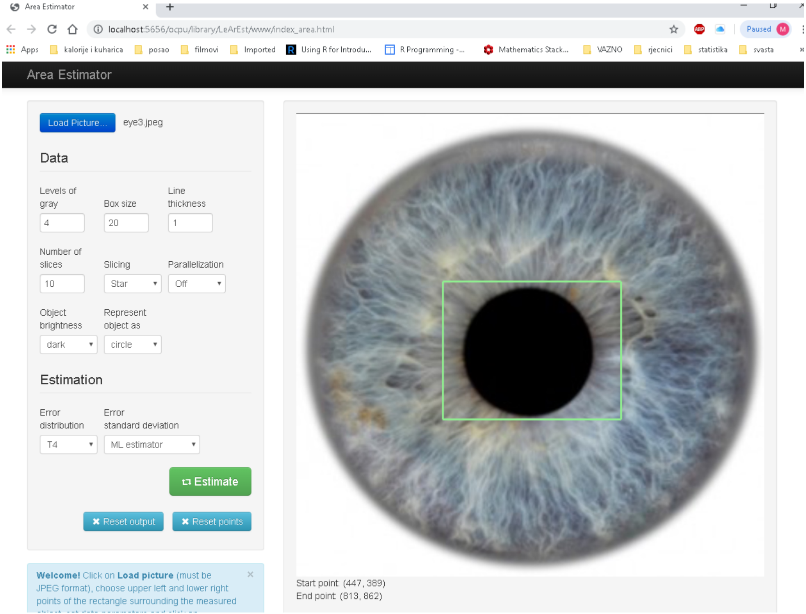
Updated R package

$https://cran.r-project.org/web/packages/LeArEst/$

Eye pupil

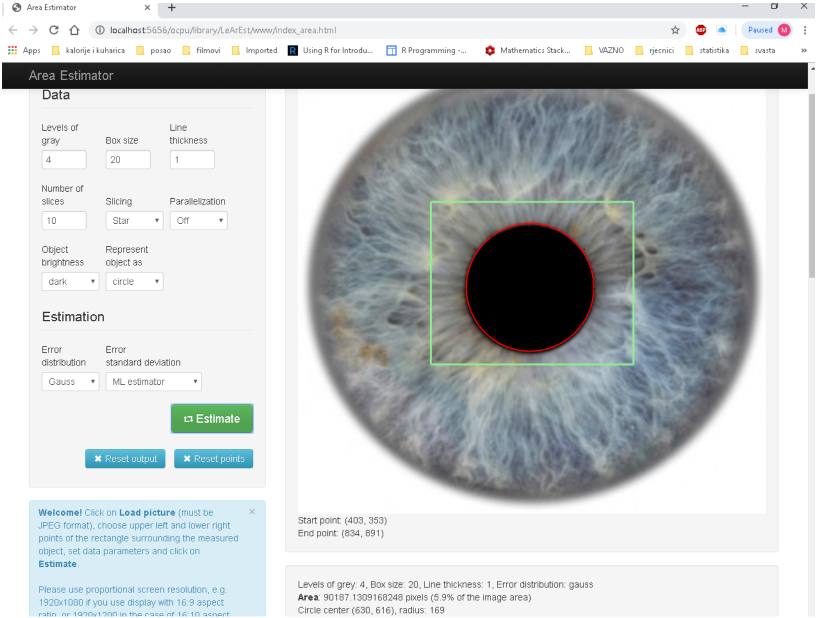

Area Estimator
localhost:5656/ocpu/library/LeArEst/www/index_area.html
Apps
kalorije i kuharica
posao
filmovi
Imported
Using R for Introdu...
R Programming -...
Mathematics Stack...
VAZNO
rjecnici
statistika
svasta
Paused
Area Estimator
Load Picture...
eye3.jpeg
Data
Levels of gray
4
Box size
20
Line thickness
1
Number of slices
10
Slicing
Star
Parallelization
Off
Object brightness
dark
Represent object as
circle
Estimation
Error distribution
T4
Error standard deviation
ML estimator
Estimate
Reset output
Reset points
Welcome! Click on Load picture (must be JPEG format), choose upper left and lower right points of the rectangle surrounding the measured object, set data parameters and click on
Levels of grey: 4, Box size: 20, Line thickness: 1, Error distribution: t4

Eye pupil

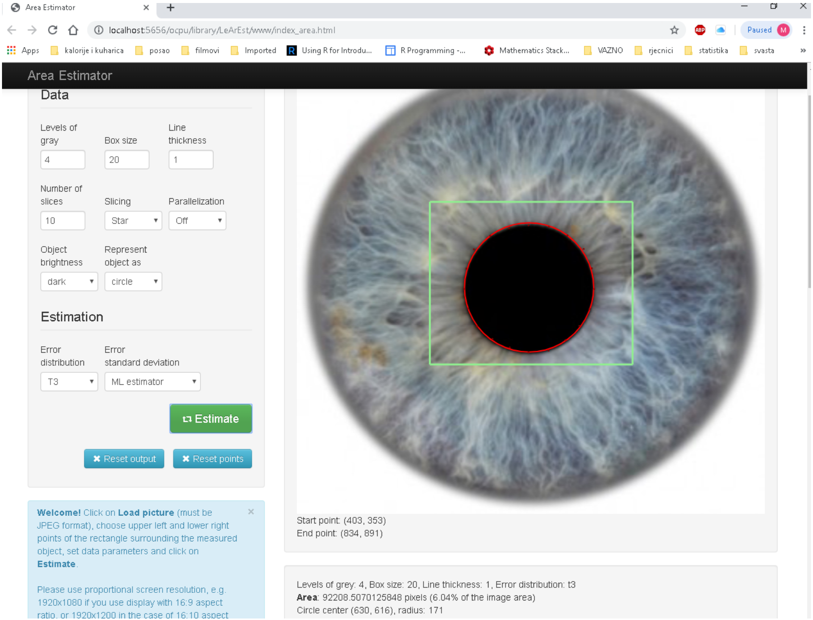


i okvir.PNG

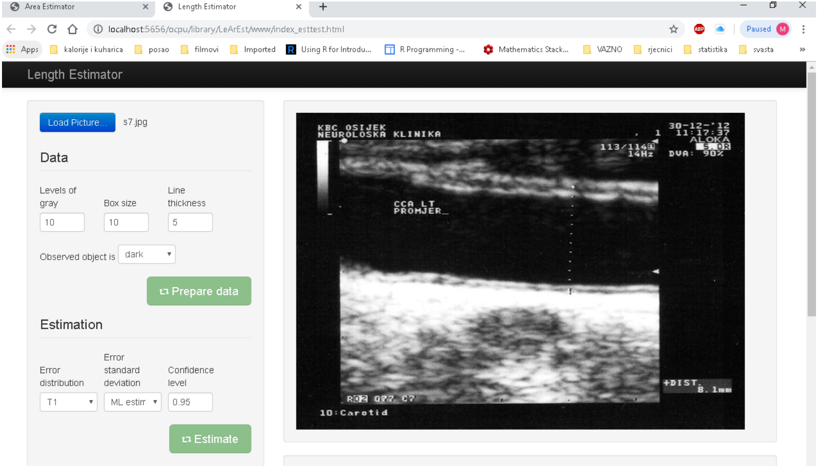
Eye pupil


Area Estimator
localhost:5656/ocpu/library/LeArEst/www/index_area.html
Paused
Apps
kalorije i kuharica
posao
filmovi
Imported
Using R for Introdu...
R Programming -...
Mathematics Stack...
VAZNO
rjecnici
statistika
svasta
Area Estimator
Data
Levels of gray
4
Box size
20
Line thickness
1
Number of slices
10
Slicing
Star
Parallelization
Off
Object brightness
dark
Represent object as
circle
Estimation
Error distribution
Gauss
Error standard deviation
ML estimator
Estimate
Reset output
Reset points
Welcome! Click on Load picture (must be JPEG format), choose upper left and lower right points of the rectangle surrounding the measured object, set data parameters and click on Estimate.
Please use proportional screen resolution, e.g. 1920x1080 if you use display with 16:9 aspect ratio, or 1920x1200 in the case of 16:10 aspect
Start point: (403, 353)
End point: (834, 891)
Levels of grey: 4, Box size: 20, Line thickness: 1, Error distribution: gauss
Area: 90187.1309168248 pixels (5.9% of the image area)
Circle center (630, 616), radius: 169

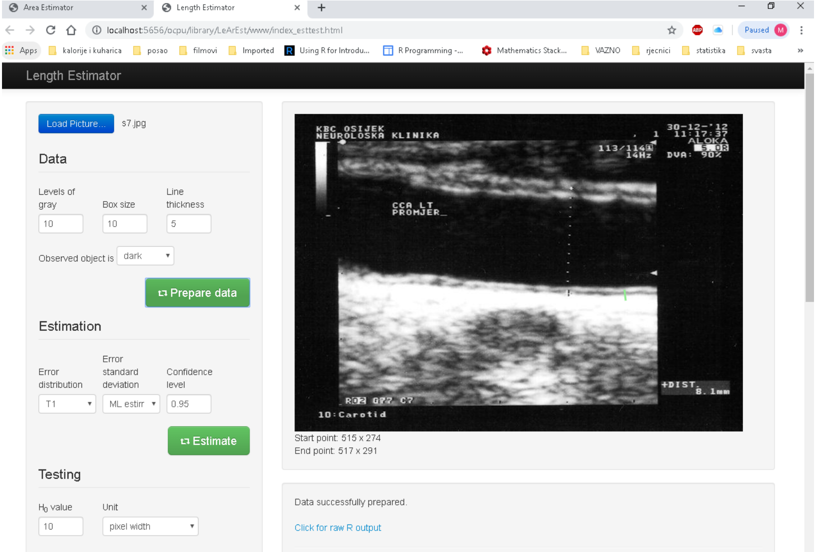
Eye pupil


Area Estimator
localhost:5656/ocpu/library/LeArEst/www/index_area.html
Apps
kalorije i kuharica
posao
filmovi
Imported
Using R for Introdu...
R Programming -...
Mathematics Stack...
VAZNO
rjecnici
statistika
svasta
Paused
Area Estimator
Data
Levels of gray
4
Box size
20
Line thickness
1
Number of slices
10
Slicing
Star
Parallelization
Off
Object brightness
dark
Represent object as
circle
Estimation
Error distribution
T3
Error standard deviation
ML estimator
Estimate
Reset output
Reset points
Welcome! Click on Load picture (must be JPEG format), choose upper left and lower right points of the rectangle surrounding the measured object, set data parameters and click on Estimate.
Please use proportional screen resolution, e.g. 1920x1080 if you use display with 16:9 aspect ratio, or 1920x1200 in the case of 16:10 aspect
Start point: (403, 353)
End point: (834, 891)
Levels of grey: 4, Box size: 20, Line thickness: 1, Error distribution: t3
Area: 92208.5070125848 pixels (6.04% of the image area)
Circle center (630, 616), radius: 171

Blood artery



Blood artery



Blood artery