Feature Extraction

degreedata.xlsx

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.1.1     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## Warning: package 'ggplot2' was built under R version 3.6.2
```

```
## Warning: package 'tibble' was built under R version 3.6.2
```

```
## Warning: package 'tidyr' was built under R version 3.6.2
```

```
## Warning: package 'purrr' was built under R version 3.6.2
```

```
## Warning: package 'dplyr' was built under R version 3.6.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
p2_degree = read_excel(path = "feature/p2-fro_degreedata.xlsx", col_names = TRUE)
p3_degree = read_excel(path = "feature/p3-fro_degreedata.xlsx", col_names = TRUE)
p4_degree = read_excel(path = "feature/p4-fro_degreedata.xlsx", col_names = TRUE)
p5_degree = read_excel(path = "feature/p5-fro_degreedata.xlsx", col_names = TRUE)
p6_degree = read_excel(path = "feature/p6-fro_degreedata.xlsx", col_names = TRUE)
p7_degree = read_excel(path = "feature/p7-fro_degreedata.xlsx", col_names = TRUE)
```

Count of branchpoints

```
# Count of branchpoints
nrow(p2_degree)
```

```
## [1] 481
```

```
nrow(p3_degree)
```

```
## [1] 910
```

```
nrow(p4_degree)
```

```
## [1] 1444
```

```
nrow(p5_degree)
```

```
## [1] 2723
```

```r
nrow(p6_degree)
```

```
## [1] 3238
```

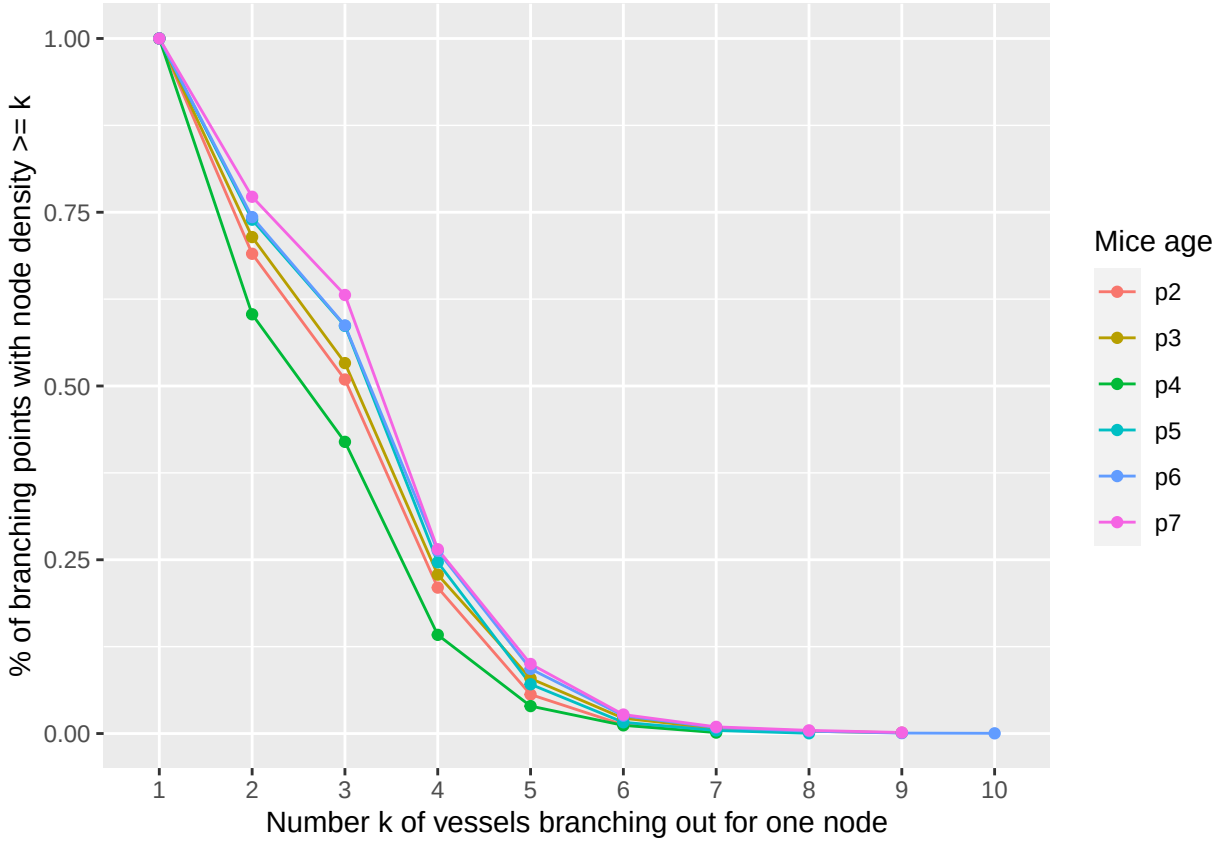
```r
nrow(p7_degree)
```

```
## [1] 3956
```

Spatial density of nodes (Branching number per node)

```r
# Branching number per node
# Spatial density of nodes
deg_spatial_density = data.frame(miceage = c(), degree = c(), density = c())

degree = p2_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p2",max(degree)), degree = 1:max(degree), density

degree = p3_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p3",max(degree)), degree = 1:max(degree), density

degree = p4_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p4",max(degree)), degree = 1:max(degree), density

degree = p5_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p5",max(degree)), degree = 1:max(degree), density

degree = p6_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p6",max(degree)), degree = 1:max(degree), density
```

```
degree = p7_degree$degree
deg_density = table(degree)[max(degree):1] %>%
  prop.table() %>%
  cumsum() %>%
  rev()
deg_spatial_density = rbind(deg_spatial_density,
                            data.frame(miceage = rep("p7",max(degree)), degree = 1:max(degree), density

deg_spatial_density$miceage = as.factor(deg_spatial_density$miceage)
deg_spatial_density$degree = as.factor(deg_spatial_density$degree)
ggplot(deg_spatial_density, aes(x = degree, y = density, color = miceage, group = miceage)) +
  geom_line() +
  geom_point() +
  labs(x = "Number k of vessels branching out for one node",
       y = "% of branching points with node density >= k") +
  scale_colour_discrete("Mice age")
```



Reference: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6076630

alldata.xlsx

```
p2_all = read_excel(path = "feature/p2-fro_alldata.xlsx", col_names = TRUE)
p3_all = read_excel(path = "feature/p3-fro_alldata.xlsx", col_names = TRUE)
p4_all = read_excel(path = "feature/p4-fro_alldata.xlsx", col_names = TRUE)
p5_all = read_excel(path = "feature/p5-fro_alldata.xlsx", col_names = TRUE)
p6_all = read_excel(path = "feature/p6-fro_alldata.xlsx", col_names = TRUE)
```

```
p7_all = read_excel(path = "feature/p7-fro_alldata.xlsx", col_names = TRUE)
```

Vessel segment - the number of vessel edges

```
nrow(p2_all)
```

```
## [1] 598
```

```
nrow(p3_all)
```

```
## [1] 1178
```

```
nrow(p4_all)
```

```
## [1] 1601
```

```
nrow(p5_all)
```

```
## [1] 3627
```

```
nrow(p6_all)
```

```
## [1] 4411
```

```
nrow(p7_all)
```

```
## [1] 5561
```

It certainly increases as it gets older.

Vessel length & Tortuosity & Width

```
mean(p2_all$length)
```

```
## [1] 11.86584
```

```
mean(p3_all$length)
```

```
## [1] 11.79226
```

```
mean(p4_all$length)
```

```
## [1] 12.09268
```

```
mean(p5_all$length)
```

```
## [1] 11.68207
```

```
mean(p6_all$length)
```

```
## [1] 11.56996
```

```
mean(p7_all$length)
```

```
## [1] 11.62666
```

```
sum(p2_all$length)
```

```
## [1] 7095.773
```

```
sum(p3_all$length)
```

```
## [1] 13891.28
```

```r
sum(p4_all$length)
```

```
## [1] 19360.39
```

```r
sum(p5_all$length)
```

```
## [1] 42370.87
```

```r
sum(p6_all$length)
```

```
## [1] 51035.11
```

```r
sum(p7_all$length)
```

```
## [1] 64655.84
```

```r
mean(p2_all$tortuosity)
```

```
## [1] 1.101782
```

```r
mean(p3_all$tortuosity)
```

```
## [1] 1.099792
```

```r
mean(p4_all$tortuosity)
```

```
## [1] 1.142496
```

```r
mean(p5_all$tortuosity)
```

```
## [1] 1.096093
```

```r
mean(p6_all$tortuosity)
```

```
## [1] 1.092839
```

```r
mean(p7_all$tortuosity)
```

```
## [1] 1.085495
```

```r
mean(p2_all$width)
```

```
## [1] 2.862876
```

```r
mean(p3_all$width)
```

```
## [1] 2.736418
```

```r
mean(p4_all$width)
```

```
## [1] 2.404747
```

```r
mean(p5_all$width)
```

```
## [1] 2.60863
```

```r
mean(p6_all$width)
```

```
## [1] 2.665949
```

```r
mean(p7_all$width)
```

```
## [1] 2.4893
```

Sum of vessel length increases as it gets older
Couldn't find any other interesting facts (differences) for length and tortuosity.
Maybe Vessel width (diameter) gets smaller as it gets older? (perform t-test?)

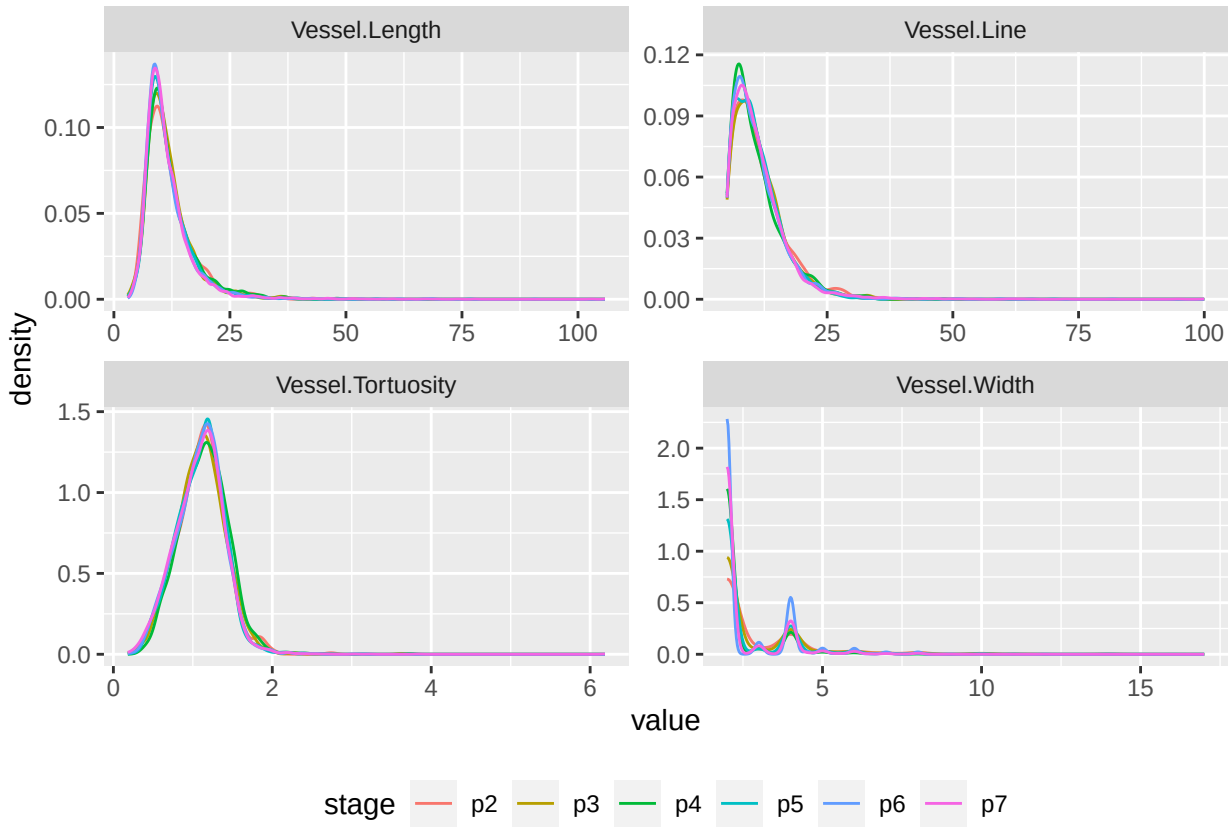
Summary

```
vessel_summary = data.frame(p2 = c(),p3 = c(),p4 = c(),p5 = c(),p6 = c(),p7 = c())
f1 = c(nrow(p2_degree),nrow(p3_degree),nrow(p4_degree),nrow(p5_degree),nrow(p6_degree),nrow(p7_degree))
f2 = c(nrow(p2_all),nrow(p3_all),nrow(p4_all),nrow(p5_all),nrow(p6_all),nrow(p7_all))
f3 = c(mean(p2_all$length),mean(p3_all$length),mean(p4_all$length),
       mean(p5_all$length),mean(p6_all$length),mean(p7_all$length))
f4 = c(mean(p2_all$tortuosity),mean(p3_all$tortuosity),mean(p4_all$tortuosity),
       mean(p5_all$tortuosity),mean(p6_all$tortuosity),mean(p7_all$tortuosity))
f5 = c(mean(p2_all$width),mean(p3_all$width),mean(p4_all$width),
       mean(p5_all$width),mean(p6_all$width),mean(p7_all$width))


m = rbind(f1,f2,f3,f4,f5)
colnames(m) = c("p2","p3","p4","p5","p6","p7")
rownames(m) = c("Number of Branchpoints", "Number of Vessel Edges","Mean of Vessel Length",
                "Mean of Vessel Tortuosity","Mean of Vessel Width")

options(digits=4)
knitr::kable(as.data.frame(m), align="c")
```

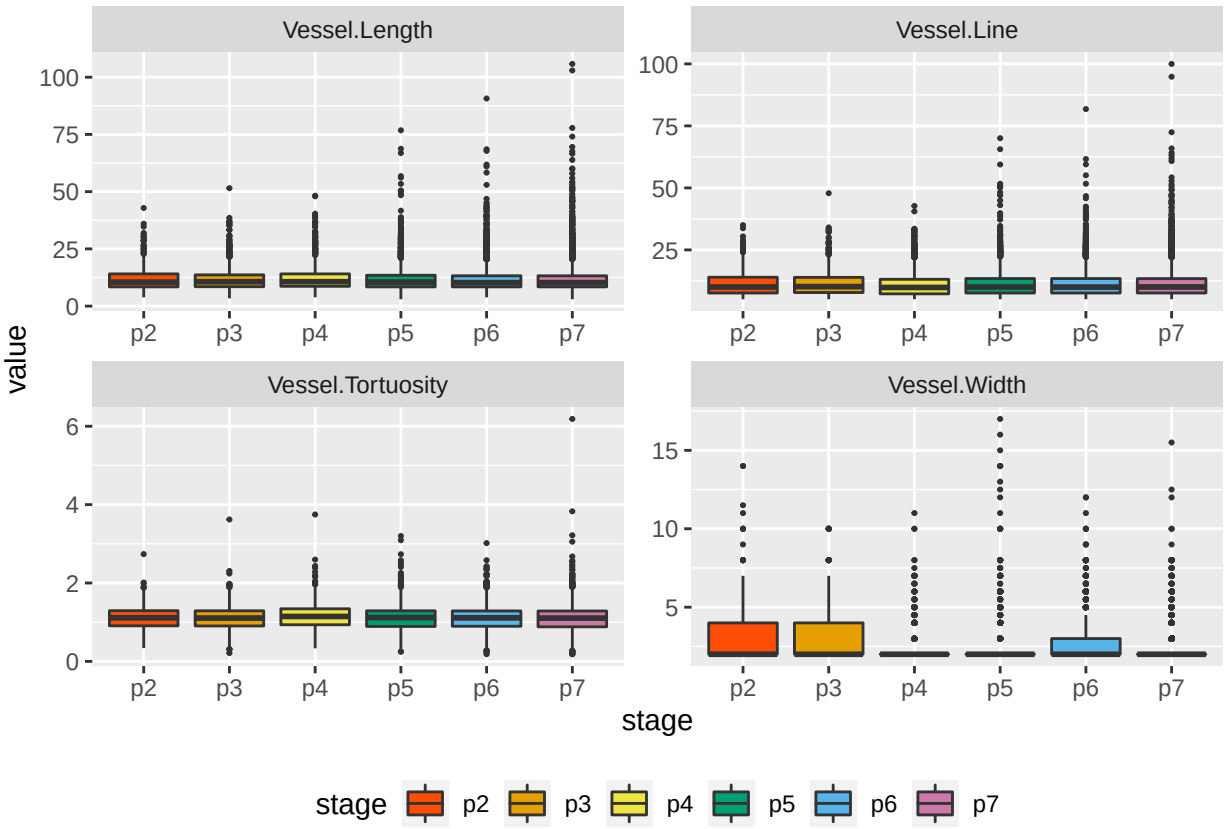
	p2	p3	p4	p5	p6	p7
Number of Branchpoints	481.000	910.000	1444.000	2723.000	3238.000	3956.000
Number of Vessel Edges	598.000	1178.000	1601.000	3627.000	4411.000	5561.000
Mean of Vessel Length	11.866	11.792	12.093	11.682	11.570	11.627
Mean of Vessel Tortuosity	1.102	1.100	1.143	1.096	1.093	1.085
Mean of Vessel Width	2.863	2.736	2.405	2.609	2.666	2.489

Density and Boxplot

```
p1to7_all <- rbind(p2_all, p3_all, p4_all, p5_all, p6_all, p7_all)
p1to7_all$stage <- c(rep("p2", nrow(p2_all)), rep("p3", nrow(p3_all)), rep("p4", nrow(p4_all)), rep("p5

colnames(p1to7_all) <- c("nodespair", "node1", "node2", "Vessel.Line", "Vessel.Length", "Vessel.Width",

p1to7_all_long <- gather(p1to7_all, key = "feature", value = "value",
                         Vessel.Line, Vessel.Length, Vessel.Width, Vessel.Tortuosity)
ggplot(p1to7_all_long, mapping = aes(x = value, fill = stage, color = stage)) + facet_wrap(.~feature, n
```




```
ggplot(p1to7_all_long, aes(x=stage, y=value, fill=stage)) + facet_wrap(.~feature, nrow = 2, scales = "fi
  geom_boxplot(width = 0.8, outlier.size=0.4) +theme(legend.position = "bottom")+  scale_x_discrete(expa
```



It seems that width variable has less variance as mouse grows.

Area

```
## Install EBImage package
### http://bioconductor.org/packages/release/bioc/html/EBImage.html
library(EBImage)
```

```
## Warning: package 'EBImage' was built under R version 3.6.1
```

```
##
## Attaching package: 'EBImage'
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
my_image2 <- readImage("../data/test_data/predictions/p2-from 5-5-2 M_prediction.tif")
my_image3 <- readImage("../data/test_data/predictions/p3-from 5-5-2 M_prediction.tif")
my_image4 <- readImage("../data/test_data/predictions/p4-from 5-5-2 M_prediction.tif")
my_image5 <- readImage("../data/test_data/predictions/p5-from 5-5-1 M_prediction.tif")
my_image6 <- readImage("../data/test_data/predictions/p6-from 5-5-2 M_prediction.tif")
my_image7 <- readImage("../data/test_data/predictions/p7-from 5-5-2 M_prediction.tif")

options(digits=4)
knitr::kable(data.frame(p2 = mean(my_image2), p3 = mean(my_image3), p4 = mean(my_image4),
                        p5 = mean(my_image5), p6 = mean(my_image6), p7 = mean(my_image7)), align = "c")
```

p2	p3	p4	p5	p6	p7
0.0185	0.0339	0.0437	0.1015	0.1193	0.1419

VAD: ratio of the total image area occupied by blood vessel area https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4901200/pdf/JBO-021-066008.pdf (page 4)

Bar plots

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:EBImage':
##
##     combine
```

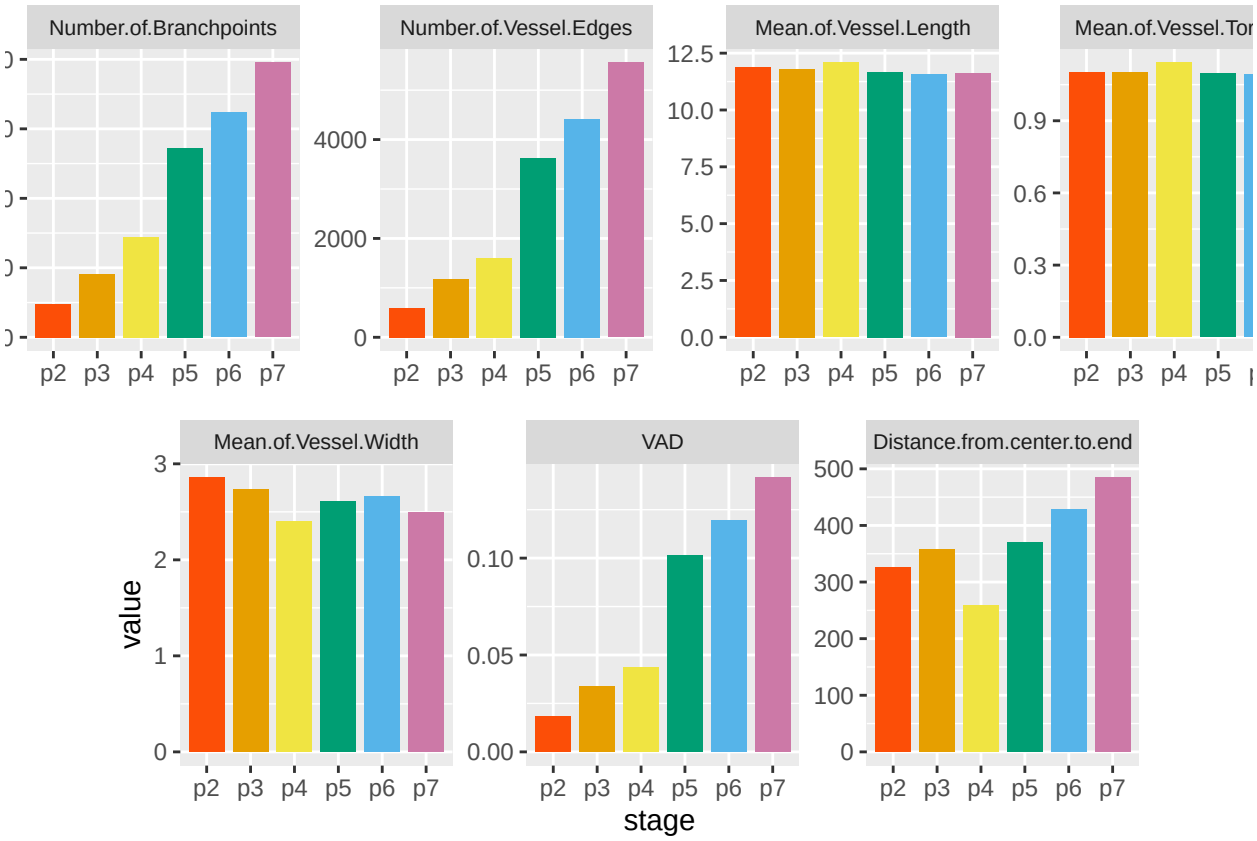
```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(cowplot)
```

```
## Warning: package 'cowplot' was built under R version 3.6.2
```

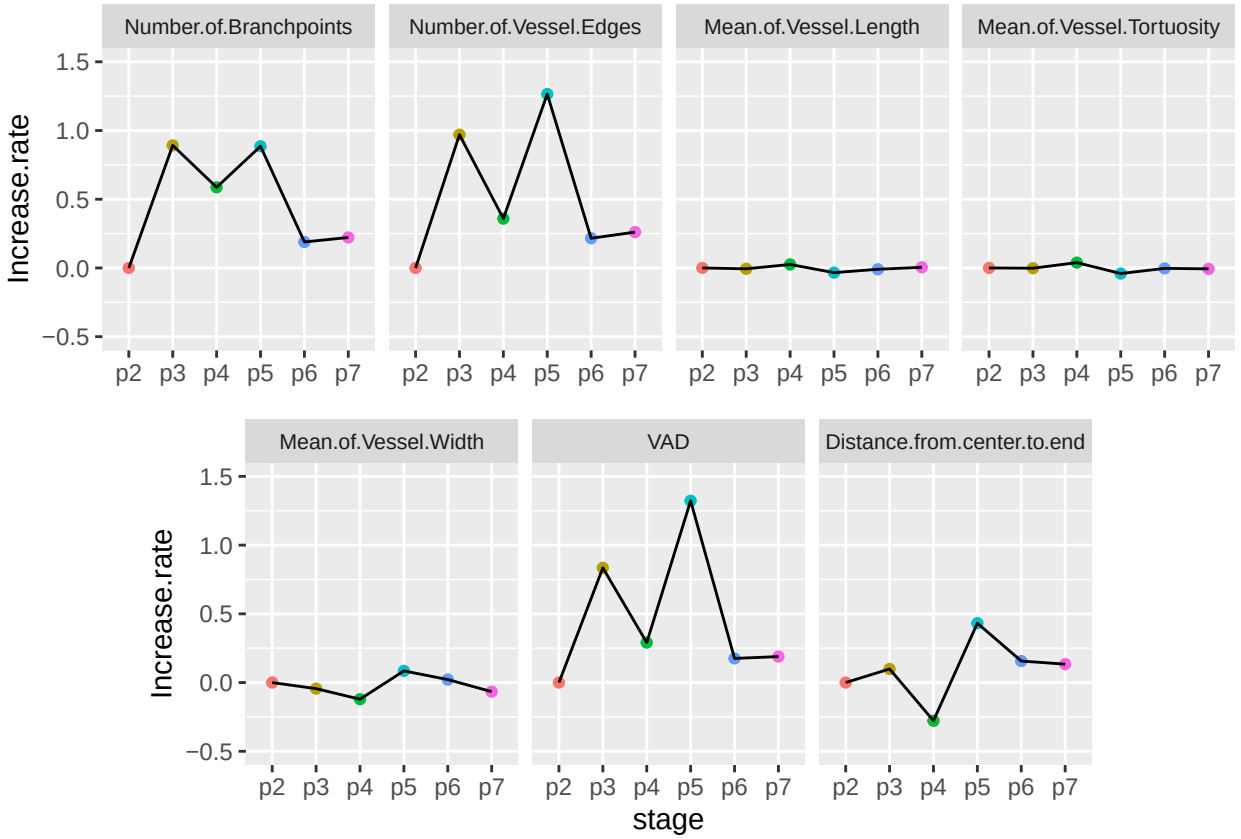
```
## VAD: ratio of the total image area occupied by blood vessel area
f6 = c(mean(my_image2), mean(my_image3), mean(my_image4),
       mean(my_image5), mean(my_image6), mean(my_image7))

##
f7 = c(max(p2_degree$distance), max(p3_degree$distance), max(p4_degree$distance),
       max(p5_degree$distance), max(p6_degree$distance), max(p7_degree$distance))

df <- data.frame(stage=c("p2","p3","p4","p5","p6","p7"),
                 f1 = f1, f2 = f2, f3 = f3, f4 = f4, f5 = f5, f6 = f6, f7 = f7)
colnames(df) <- c("stage", "Number.of.Branchpoints", "Number.of.Vessel.Edges", "Mean.of.Vessel.Length",
                  "Mean.of.Vessel.Tortuosity", "Mean.of.Vessel.Width", "VAD", "Distance.from.center.to.
df_long <- gather(df, key = "feature", value = "value",
                  Number.of.Branchpoints, Number.of.Vessel.Edges, Mean.of.Vessel.Length,
                  Mean.of.Vessel.Tortuosity, Mean.of.Vessel.Width, VAD, Distance.from.center.to.end)

df_long$feature <- factor(df_long$feature, levels = c("Number.of.Branchpoints",
                                                      "Number.of.Vessel.Edges",
                                                      "Mean.of.Vessel.Length",
                                                      "Mean.of.Vessel.Tortuosity",
                                                      "Mean.of.Vessel.Width",
                                                      "VAD",
                                                      "Distance.from.center.to.end"))

feature.p <- ggplot(data=df_long, aes(x=stage, y=value, fill = stage)) +
  geom_bar(stat="identity", width=0.8, show.legend=FALSE)  + facet_wrap(.~feature, nrow = 1, scales = "


library(egg)
library(gridExtra)
```

```
feature.p1 <- feature.p  %+% df_long[1:24,] + labs(x = NULL)
feature.p2 <- feature.p  %+% df_long[-(1:24),]
grid.arrange(grobs = lapply(
  list(feature.p1, feature.p2),
  set_panel_size,
  width = unit(3.6, "cm"),
  height = unit(4, "cm")
))
```




```
df.rate <- data.frame(stage=c("p2","p3","p4","p5","p6","p7"),
                      f1 = c(0,f1[2:6]/f1[1:5]-1), f2 = c(0,f2[2:6]/f2[1:5]-1),
                      f3 = c(0,f3[2:6]/f3[1:5]-1), f4 = c(0,f4[2:6]/f4[1:5]-1),
                      f5 = c(0,f5[2:6]/f5[1:5]-1), f6 = c(0,f6[2:6]/f6[1:5]-1),
                      f7 = c(0,f7[2:6]/f7[1:5]-1))

colnames(df.rate) <- c("stage", "Number.of.Branchpoints", "Number.of.Vessel.Edges", "Mean.of.Vessel.Leng
                       "Mean.of.Vessel.Tortuosity", "Mean.of.Vessel.Width", "VAD", "Distance.from.center.to.

df.rate_long <- gather(df.rate, key = "feature", value = "Increase.rate",
                       Number.of.Branchpoints, Number.of.Vessel.Edges, Mean.of.Vessel.Length,
                       Mean.of.Vessel.Tortuosity, Mean.of.Vessel.Width, VAD, Distance.from.center.to.end)

df.rate_long$feature <- factor(df.rate_long$feature, levels = c("Number.of.Branchpoints",
                                                                "Number.of.Vessel.Edges",
                                                                "Mean.of.Vessel.Length",
                                                                "Mean.of.Vessel.Tortuosity",
                                                                "Mean.of.Vessel.Width",
```

```
                                                    "VAD",
                                                    "Distance.from.center.to.end"))

rate.p <- ggplot(data=df.rate_long, aes(x=stage, y=Increase.rate, group = 1)) +
  geom_point(aes(colour = factor(stage))) + geom_line() + facet_wrap(.~feature, nrow = 1) + theme(strip


library(egg)
library(gridExtra)
rate.p1 <- rate.p  %+% df.rate_long[1:24,] + labs(x = NULL)
rate.p2 <- rate.p  %+% df.rate_long[-(1:24),]
grid.arrange(grobs = lapply(
  list(rate.p1, rate.p2),
  set_panel_size,
  width = unit(3.6, "cm"),
  height = unit(4, "cm")
))
```



======= >>>>>>> dd966e4653d2f0f66ce62b7246cceebe2b712041