BELT Data Definitions and their respective column in Excel

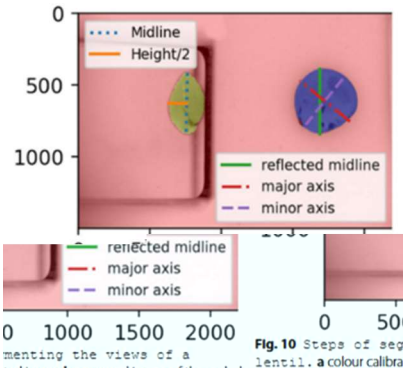
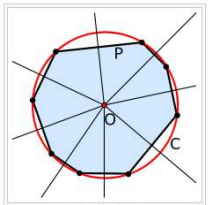


Fig. 10 Steps of segmenting the views of a lentil. a colour calibra

F. Height(mm): height from midline of lentil x2
G. Area $(mm)^{2:}$ area of the lentil from the above view
H. Perimeter (mm): The measurement in mm around a seed
I. Major axis (mm): longest width of seed
J. Minor Axis (mm): narrowest width of seed
K. Equivalent Diameter (mm): length of the reflected midline
L. Roundness= (4*area/(pi*major_axis^2): measurement based on the major axis (2d top view) https://doi.org/10.1186/s13007-020-00591-8
    a. Ratio = area of the seed: area of the circumscribed circle (The circle that themajor axis creates
        i. Higher the value the more the seed matches the reference circle
        ii. 

Circumscribed circle, C, and circumcenter, O, of a cyclic polygon, P

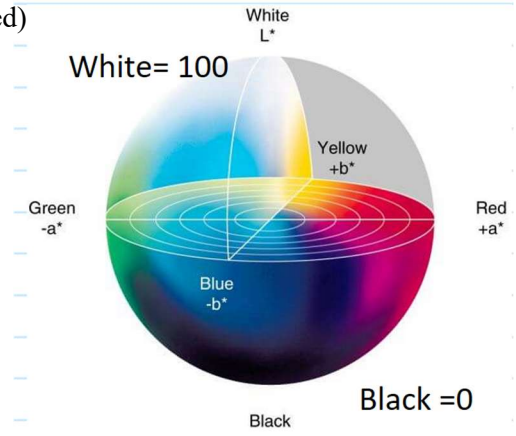
https://en.wikipedia.org/wiki/Circumscribed_circle

M. Circularity= (4*pi*area/(perimeter^2)) : measurement based on the Perimeter (2d top view)
    a. Ratio = area: perimeter: normal sphere
    b. Common in image processing; 1=1 circle. If below 1 it is noncircular shape
N. Volume (mm^3): calculated volume using an ellipsoid equation (major and minor and height as the axes)
O. Surface area (mm^2): rough avg surface area using the volume
P. Sphericity =(pi^(1/3) *(6*volume)^(2.3))/(surface area): measurement based on volume and surface area
    a. Plumpness: how filled the seed is
Q. Number pixels per component:
    a. Sample points (or # of pixels per seed)
    b. An entire picture including the background contains ~3,300,000 pixels
        i. The average pixel value was ~170,868

i. 
ii. This is the size proportion of the # of pixel compared to the pixel size of the camera (meant to show that the number of pixels represents a seed)

Colour Component data

1. L measurements (l)
    R. Mean : Mean Darkness/lightness of seed (avg darkness value of the seed)
    S. STD: the standard deviation from pixel to pixel
    T. Min: the darkest pixel on seed
    U. 25: 25% more than the min value (25% more white)
    V. 50: 50% more than the min value (50% more white)
    W. 75: 75% more than the min value (75% more white)
    X. Max: the lightest pixel on seed
2. (red+ to green-; measurement): (a)
    Y. Mean: mean value of red to green of seed (avg red-green hue of seed)



https://www.nixsensor.com/blog/measure-color-accuracy/

Z. STD: the deviation from pixel to pixel on seed

AA. Min (a_min) : the greenest pixel of seed

AB. a_25: 25% more than the min value (25% more red)

AC. a_50: 50% more than the min value (50% more red)

AD. a_75: 75% more than the min value (75% more red)

AE. Max: the reddest pixel of seed

3. (Yellow+ to blue-; measurement): (b)

AF. Mean (b_mean) : Mean value of yellow to blue of seed (avg yellow-blue hue of seed)

AG. STD (b_std): the deviation from pixel to pixel on seed

AH. Min (b_min) : bluest pixel of seed

AI. b_25: 25% more than the min value (25% more yellow)

AJ. b_50: 50% more than the min value (50% more yellow

AK. b_75: 75% more than the min value (75% more yellow)

AL. Max (b_max) : yellowest pixel of seed

Cluster Data

AM. **Cluster distance: either the distances between cluster or distance between the centers of clusters**

- Uncertain about this.

AN. Dark cluster-l: The lightness value of the dark cluster (would expect lower L value)

AO. Dark cluster-a: Red to green value of the dark cluster (how red or green it is)

AP. Dark cluster-b yellow to blue value of the dark cluster (how yellow or blue it is)

AQ. **Dark cluster proportion: proportion of Dark cluster on seed to light clusters**

AR. Light cluster-l: the lightness value of the light cluster (would expect higher L value)

AS. Light cluster-a: Red to green value of light cluster (how red to green the cluster is)

AT. Light cluster-b: yellow to blue value of light cluster (how yellow to blue the cluster is)

AU. **Light cluster proportion: proportion of light clusters on seed to dark clusters**

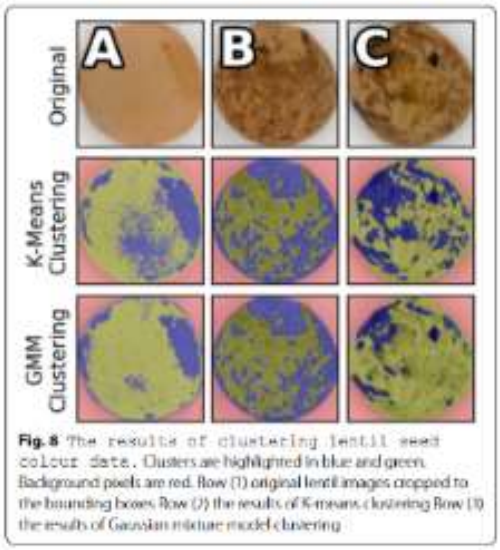


Fig. 8 The results of clustering lentil seed colour data. Clusters are highlighted in blue and green. Background pixels are red. Row (1) original lentil images cropped to the bounding boxes Row (2) the results of K-means clustering Row (3) the results of Gaussian mixture model clustering

Table 6 Summary of Euclidean distances between clusters, self-reported scores and cluster populations when L*a*b* colour pixels are clustered

Clustering metric	Example A	Example B	Example C
K-means cluster			
Distance	3.58	11.75	13.32
GMM cluster			
Distance	1.03	10.27	13.09
GMM Average			
Log likelihood	− 6.90	− 8.50	− 8.89
K-means cluster population			
(Green/Blue)	64% / 36%	51% / 49%	59% / 41%
GMM cluster population			
(Green/Blue)	73% / 27%	53% / 47%	83% / 17%

Example labels refer to Fig. 8. The cluster distances were calculated via ΔE^* CIEDE2000. GMM: Gaussian mixture model

cluster centres for example A was below JND but the K-means colour distance was not. K-means clustering was ultimately not appropriate for this application as the

https://doi.org/10.1186/s13007-020-00591-8