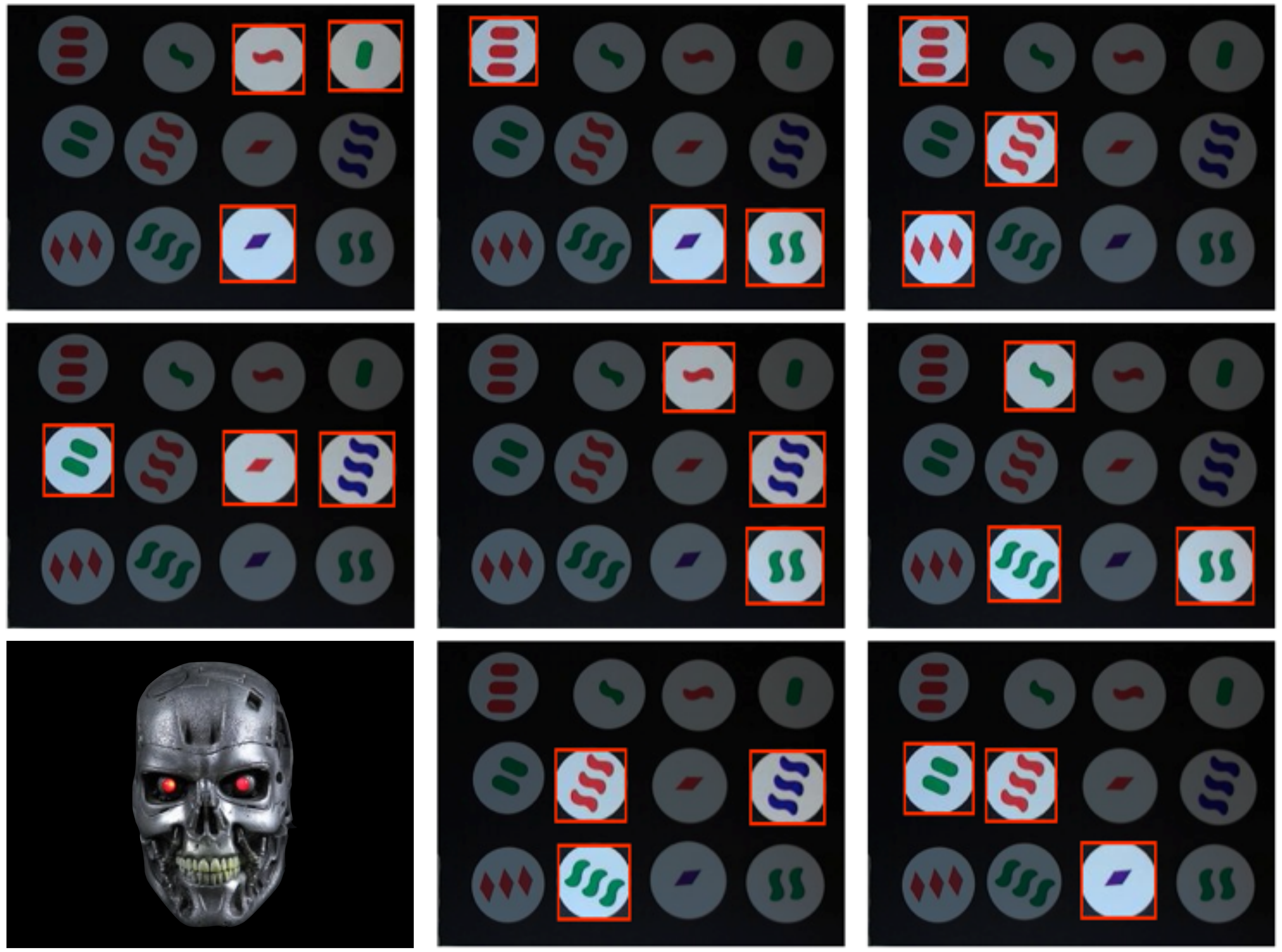
A program to play Set



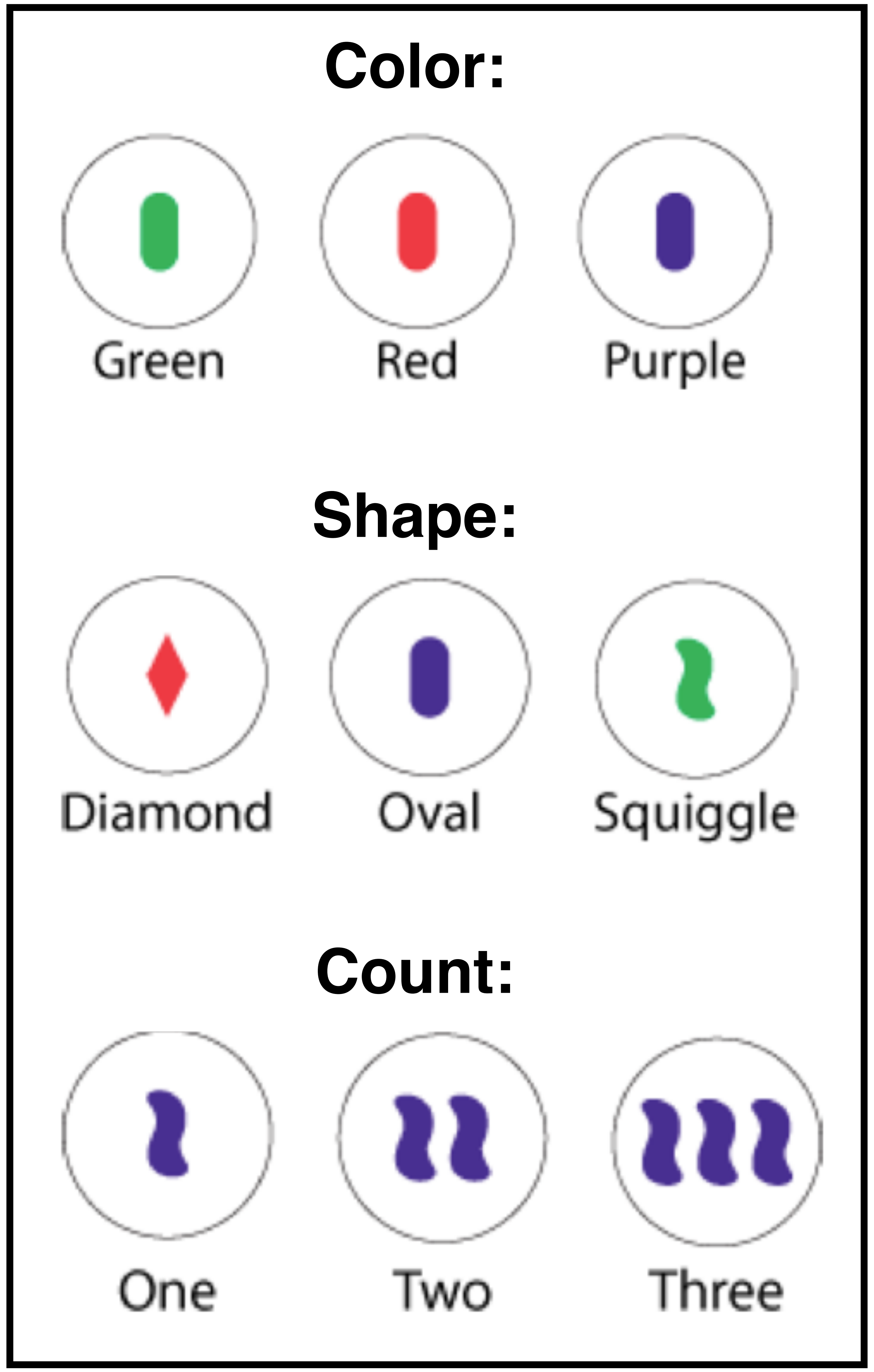
Mathematica plays the Set card game over a webcam.

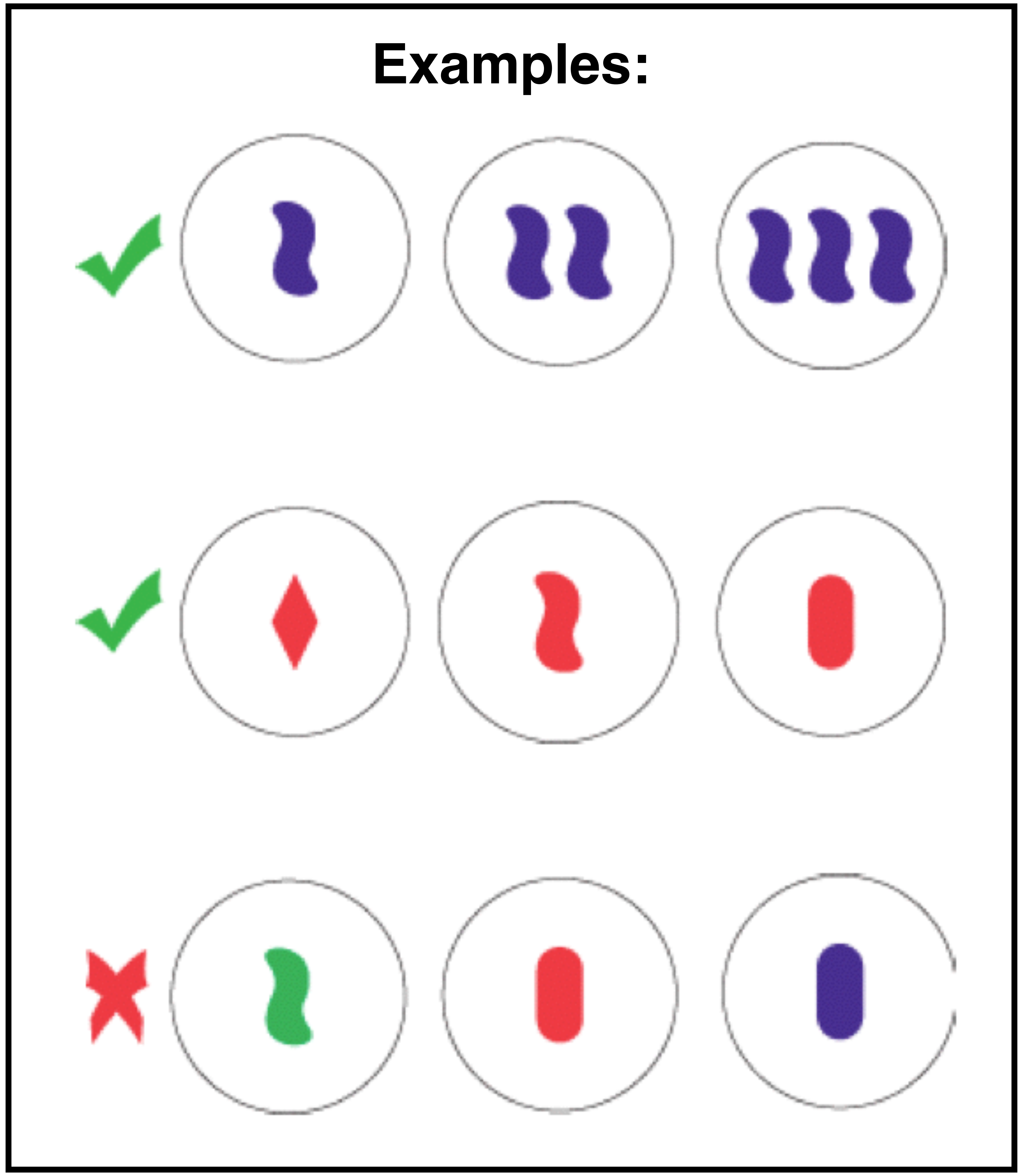
Appfolio Tech Talk
Justin Pearson
Thursday, Jan 11, 2018 11-12pm
Engineering Square and G2W

Outline

- Set
- Mathematica
- Parsing the image
    - Count
    - Color
    - Shape
- Live demo omg

For each of 3 categories: all same OR all different





Outline

- Set

- Mathematica
- Parsing the image
  - Count
  - Color
  - Shape
- Live demo omg

mma-elevator-pitch.nb

In[1]:= $\mathtt{Solve}\left[a x^2 + b x + c == 0, x\right]$

Out[1]= $\left\{\left\{x \to \frac{-b - \sqrt{b^2 - 4 a c}}{2 a}\right\}, \left\{x \to \frac{-b + \sqrt{b^2 - 4 a c}}{2 a}\right\}\right\}$

In[2]:= DSolve[{
x''[t] + 3 x'[t] + 2 x[t] == 0,
x[0] == 1, x'[0] == 0}, x[t], t]

Out[2]= $\left\{\left\{x[t] \to e^{-2 t} \left(-1 + 2 e^{t}\right)\right\}\right\}$

In[3]:= $\int_{-\infty}^{\infty} \frac{\mathtt{Sin}[x]}{x} \, dx$

Out[3]= π

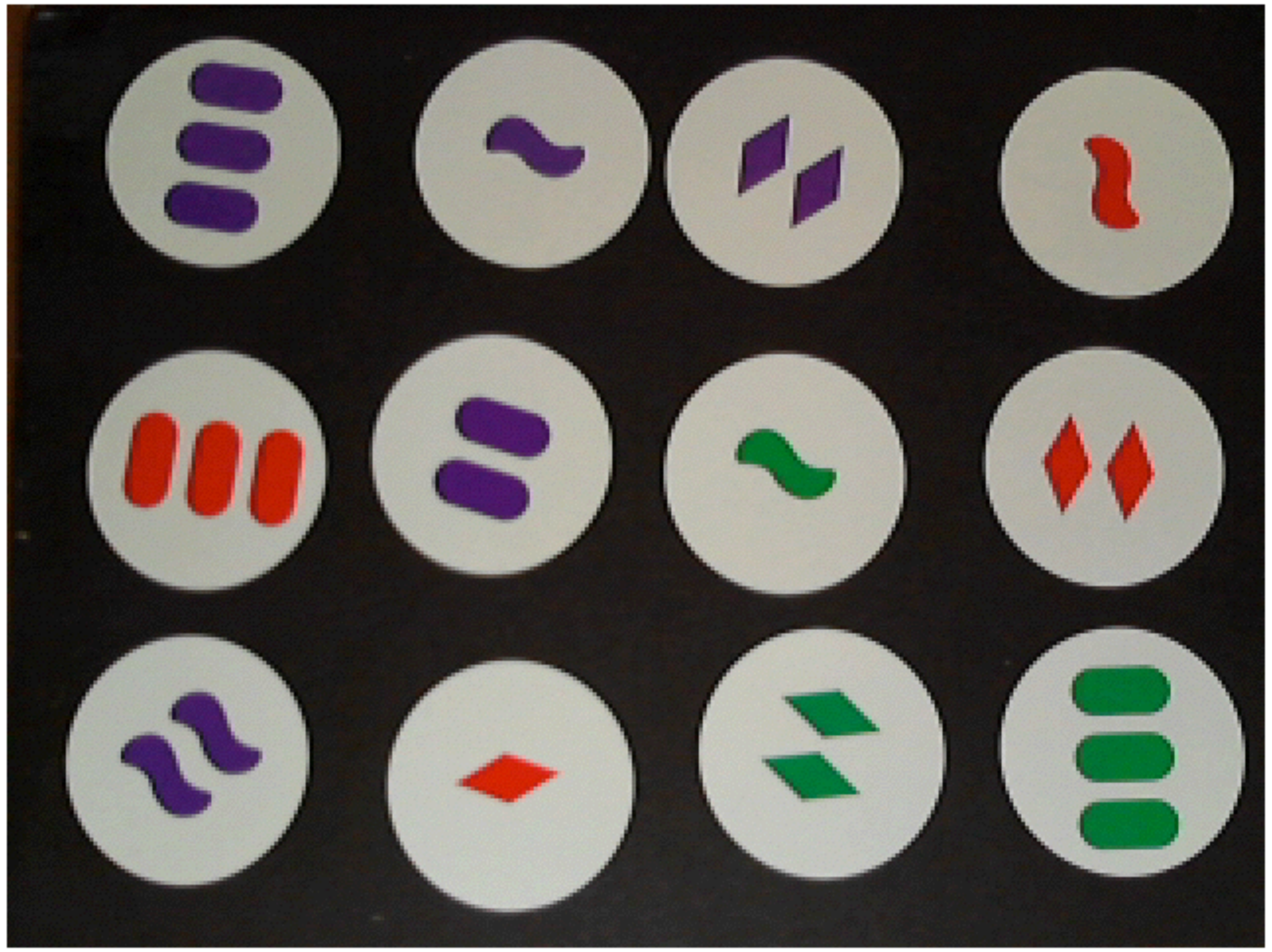
Outline

- Set
- Mathematica

- Parsing the image
    - Count
    - Color
    - Shape
- Live demo omg

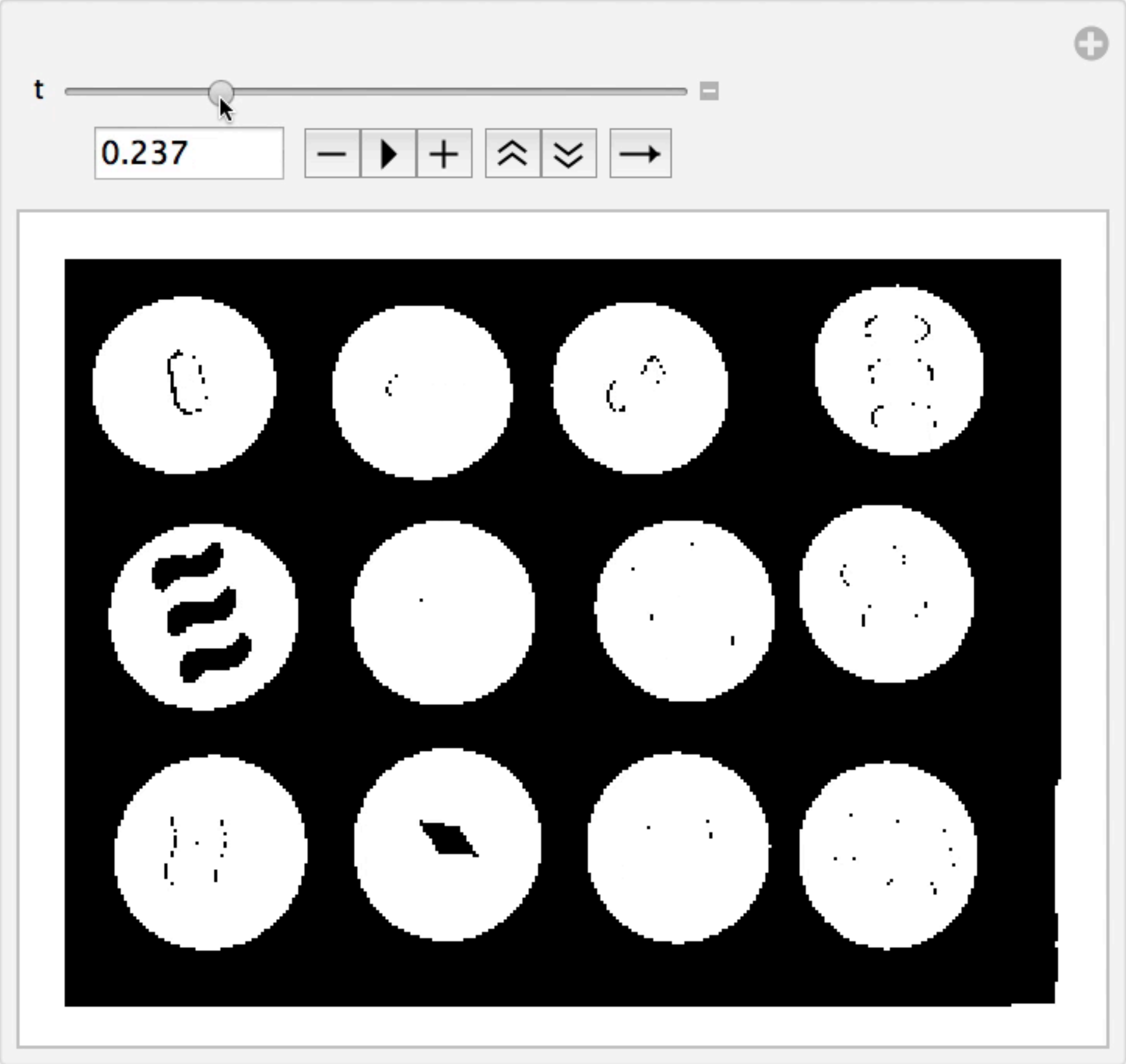
```
In[1]:= $ImagingDevices

Out[1]= {FaceTime HD Camera (Built-in), Logitech Camera}

In[2]:= $ImagingDevice = "Logitech Camera"

Out[2]= Logitech Camera

In[3]:= CurrentImage[]
```

Out[3]=



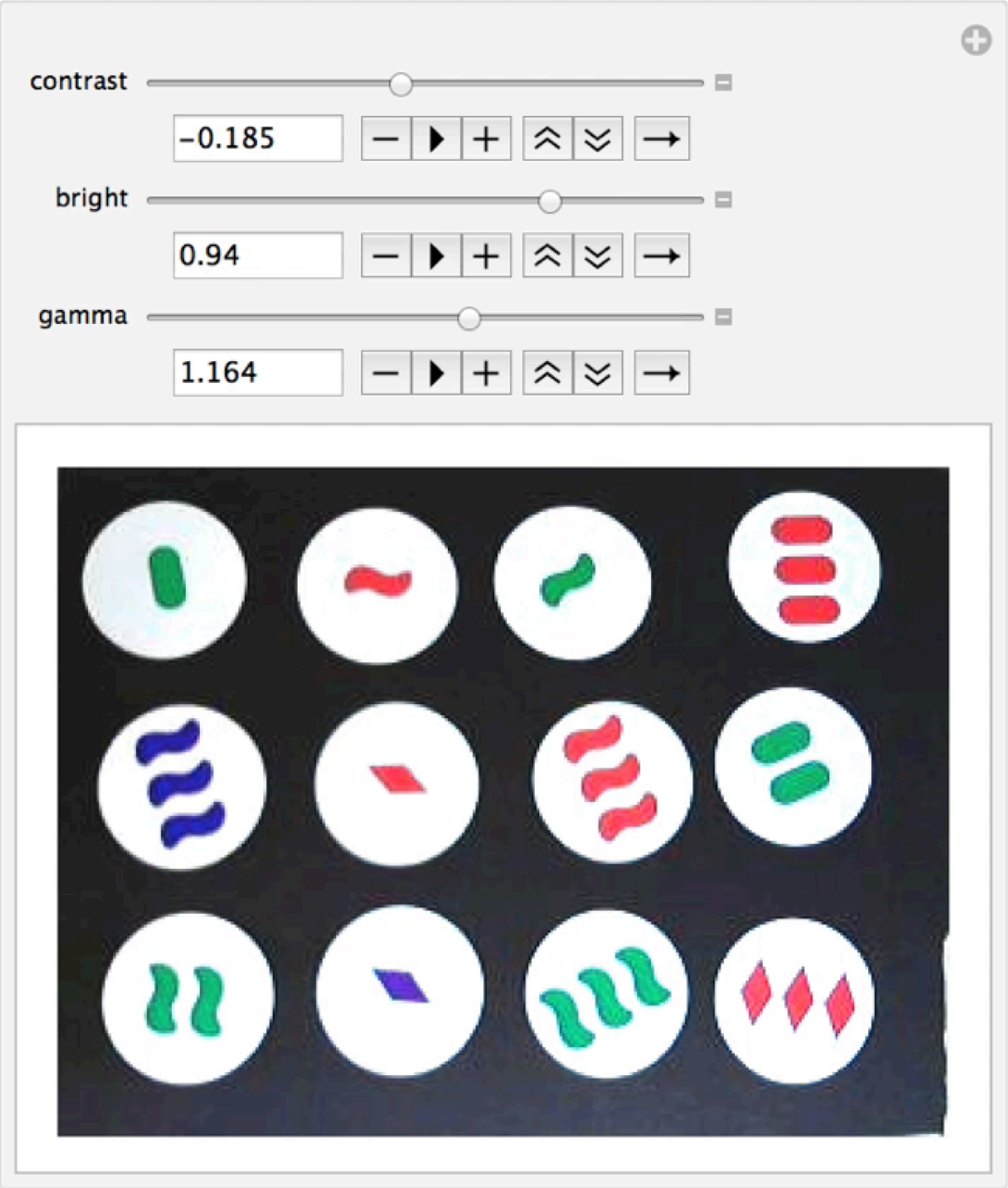
In[28]:=
```
Manipulate[
 Binarize[frame, t],
 {{t, .5}, 0, 1, Appearance → "Open"}
]
```

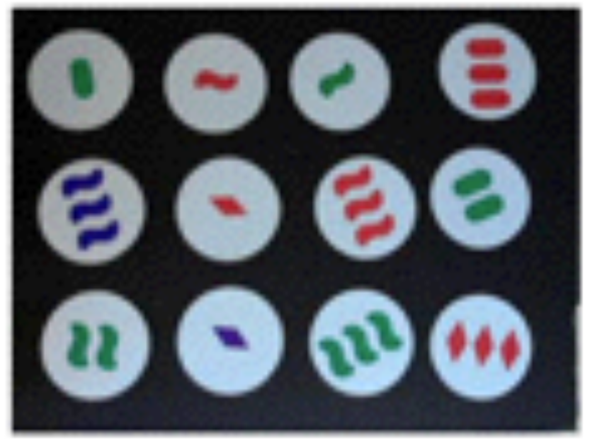
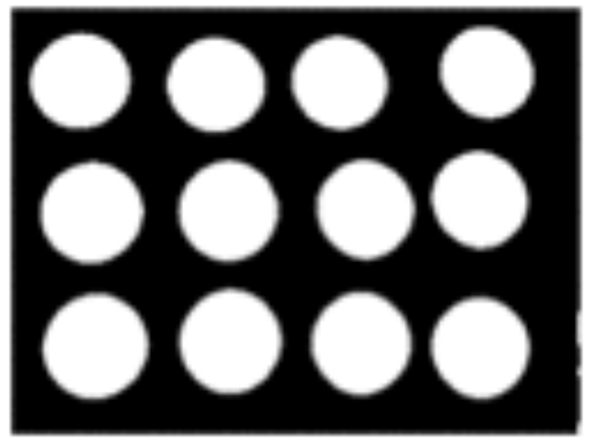
t[28]=



In[25]:=
```
Manipulate[
 ImageAdjust[frame, {contrast, bright, gamma}],
 {{contrast, .2}, -2, 2, Appearance → "Open"},
 {{bright, .6}, -2, 2, Appearance → "Open"},
 {{gamma, 1}, 0, 2, Appearance → "Open"}
]
```

Out[25]=

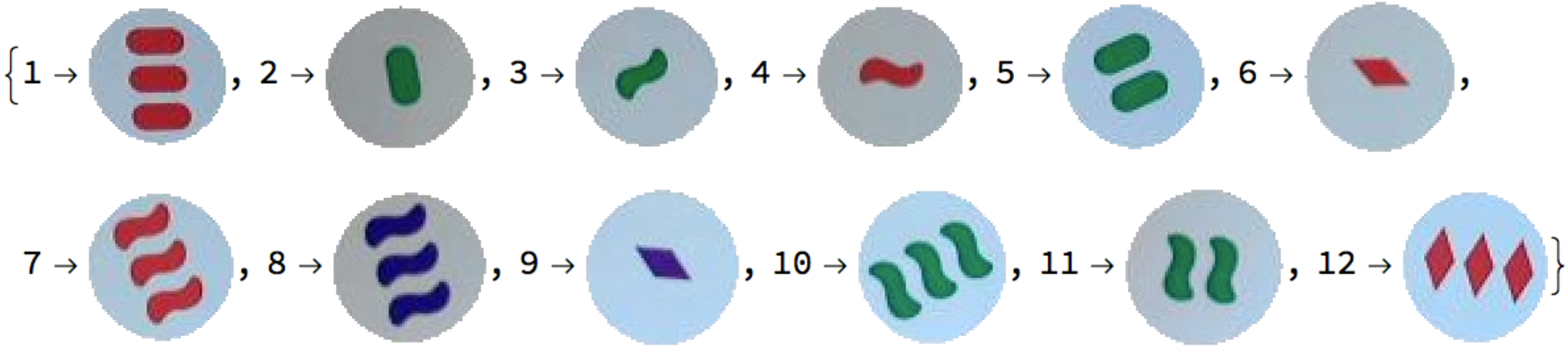


Frame to cards

In[80]:=
```
ComponentMeasurements[{frame, mask // MorphologicalComponents}, "MaskedImage"]
```

Out[80]=



Card to blobs

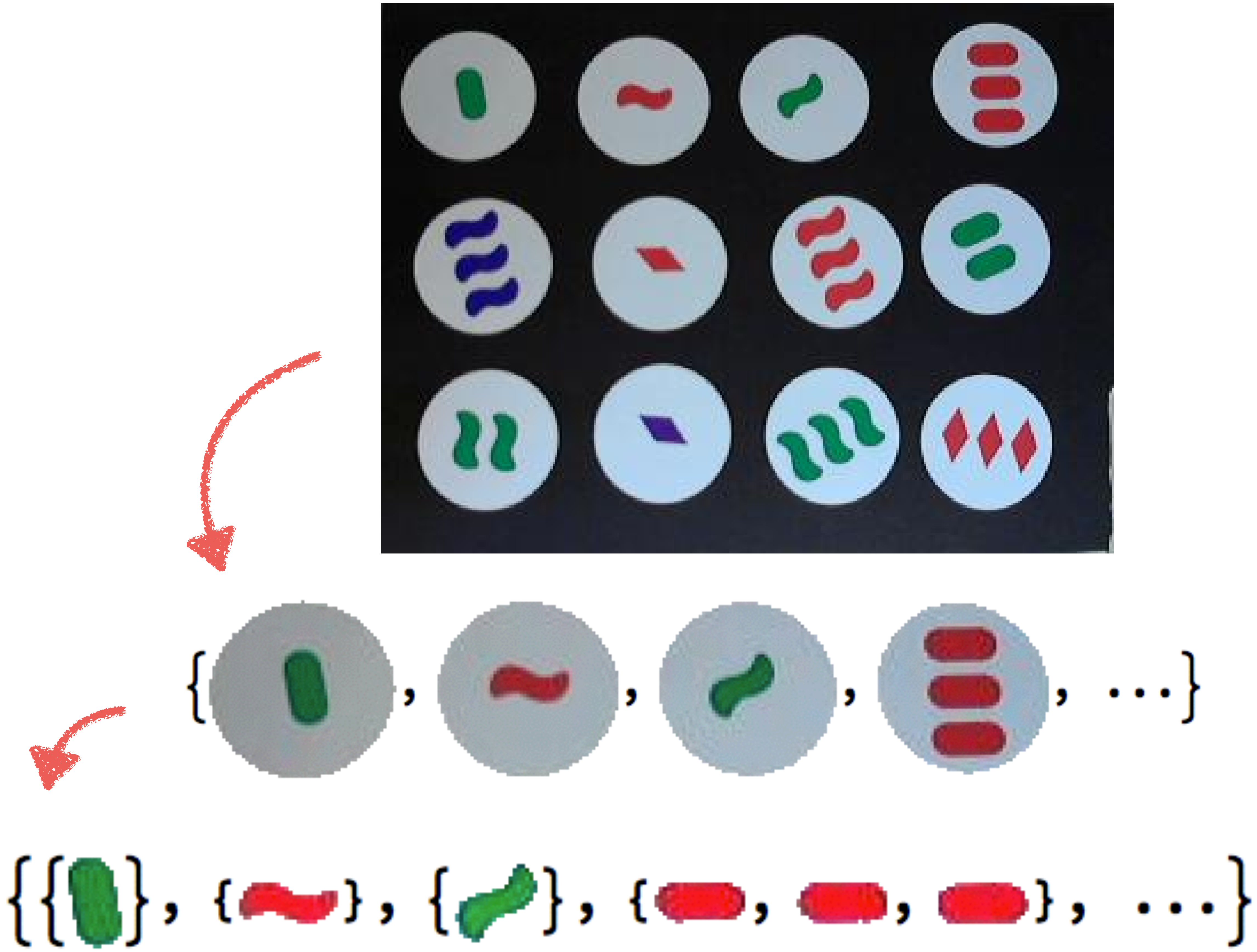
```
In[68]:= ComponentMeasurements[
  {card, mask // MorphologicalComponents},
  "MaskedImage"]
```

Out[68]=

{1 → , 2 → , 3 → }

Outline

- Set
- Mathematica
- Parsing the image 100ms
  - Count
  - Color
  - Shape
- Live demo omg



```
{
 {1, Green, "oval"},
     {1, Red, "squiggly"},
         {1, Green, "squiggly"},
             {3, Red, "oval"},
                  ...
}
```

In[134]:=

```
colors = ComponentMeasurements[
  {card, mask // MorphologicalComponents},
  {"MaskedImage", "Median"} ]
```

median RGB color of blob

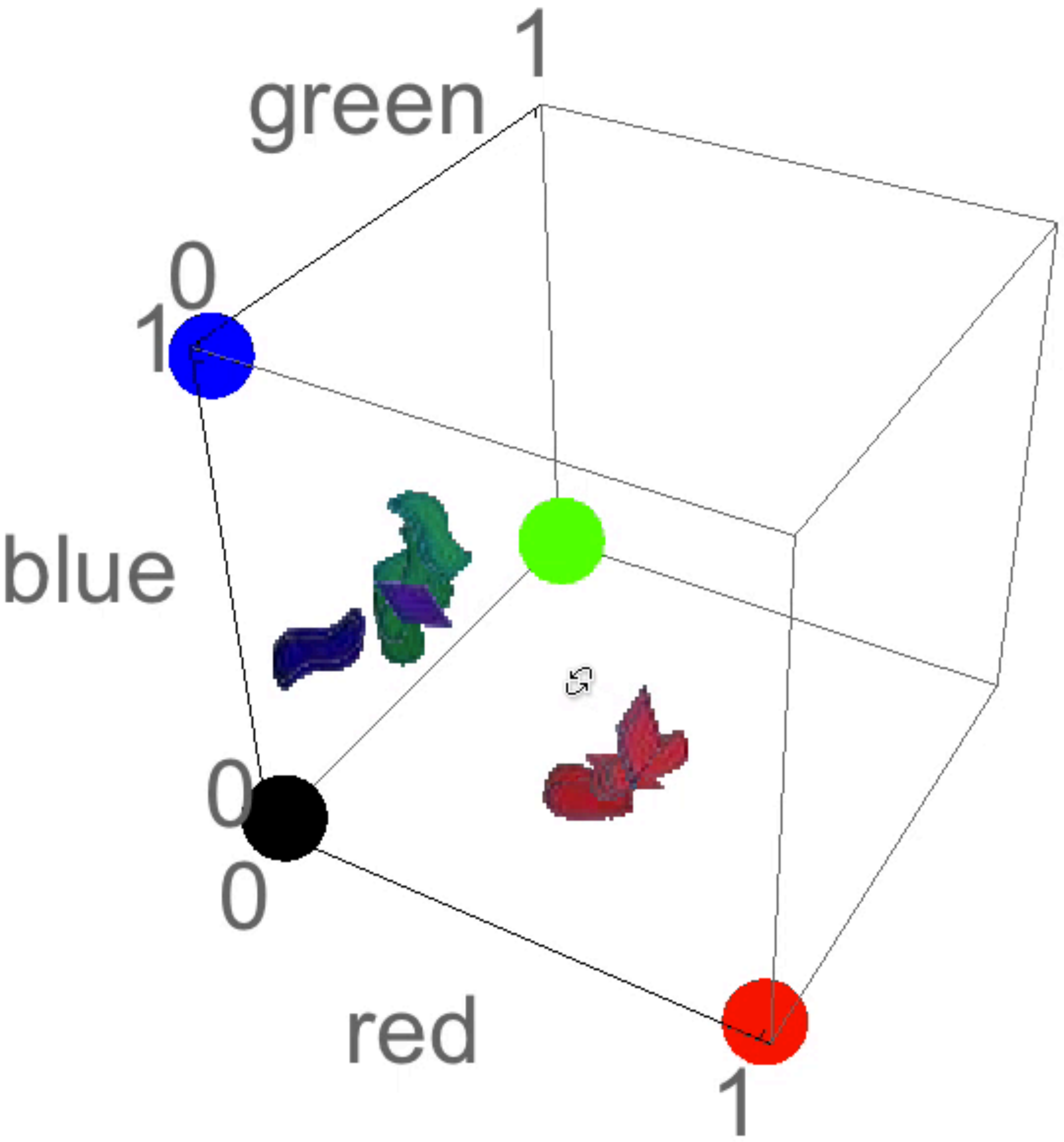
Out[134]=

{1 → {, {0.631373, 0.203922, 0.258824, 1.}},

2 → {, {0.635294, 0.211765, 0.262745, 1.}},

3 → {, {0.643137, 0.219608, 0.278431, 1.}}}

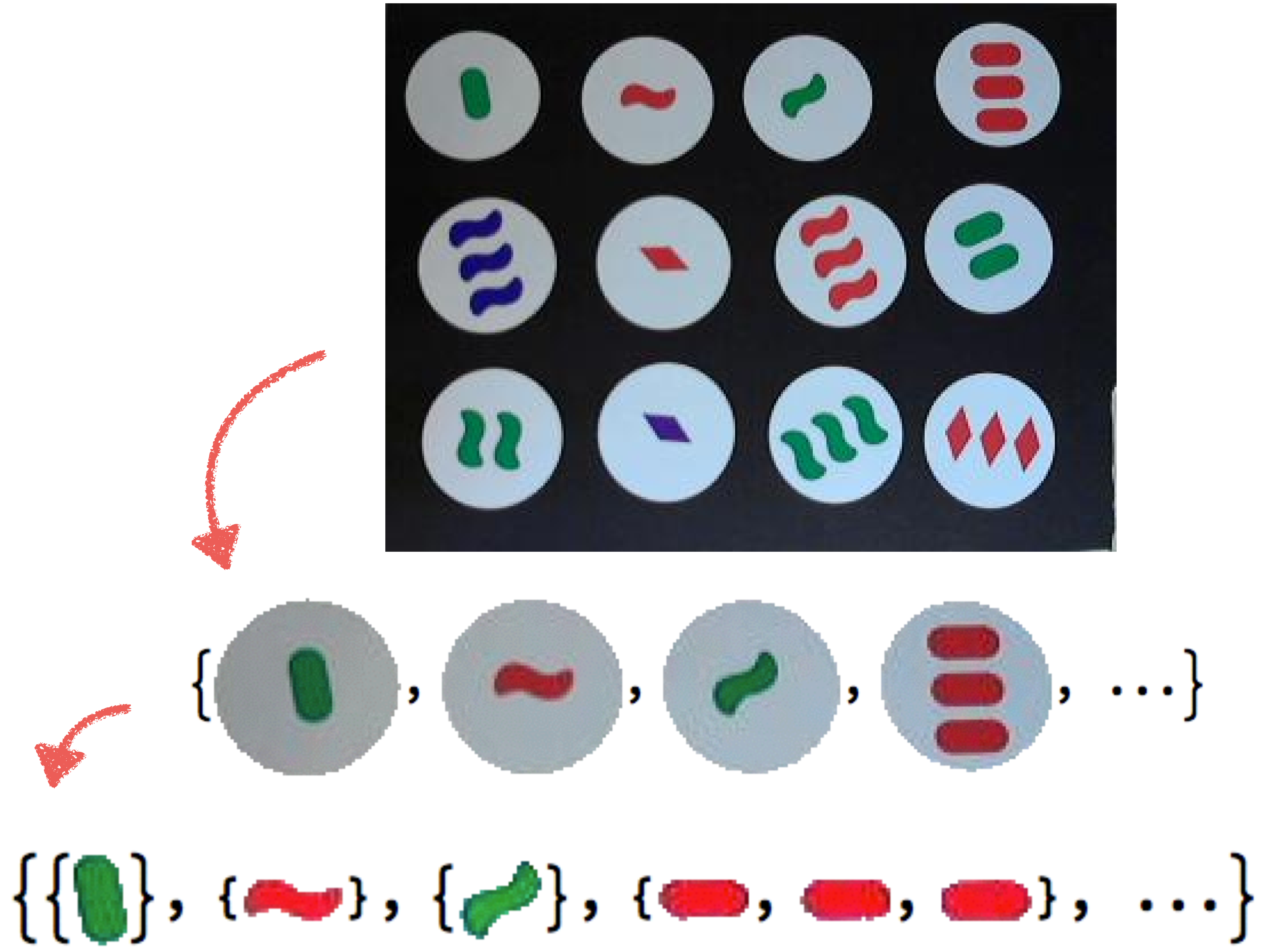
red green blue alpha



green
1
0
1
blue
0
0
red
1

Outline

- Set
- Mathematica
- Parsing the image ✓
  - Count ✓
  - Color ✓
  - Shape
- Live demo omg



```
{
 {1, Green, "oval"},
     {1, Red, "squiggly"},
         {1, Green, "squiggly"},
             {3, Red, "oval"},
                  ...
}
```

Shape classification

labeled data set

$\{ (v_1,lab_1), (v_2,lab_2), \ldots. \}$

trains

feature extractor

$$\begin{bmatrix}7\\6\\3\\ \vdots \\4\end{bmatrix} \in \mathbb{R}^n \xrightarrow{\text{classifier}} \text{“diamond”} \in \{\text{diamond}, \text{oval}, \text{squiggly}\}$$

feature vector

label

feature selection

In[202]:=

ComponentMeasurements[

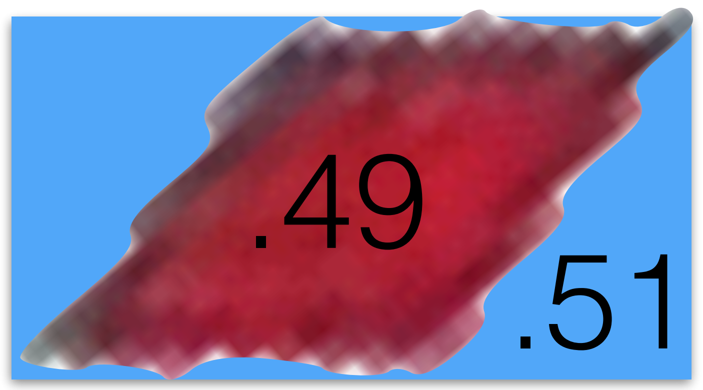
{,  // MorphologicalComponents},

{"MaskedImage", "Rectangularity", "ConvexCoverage"},

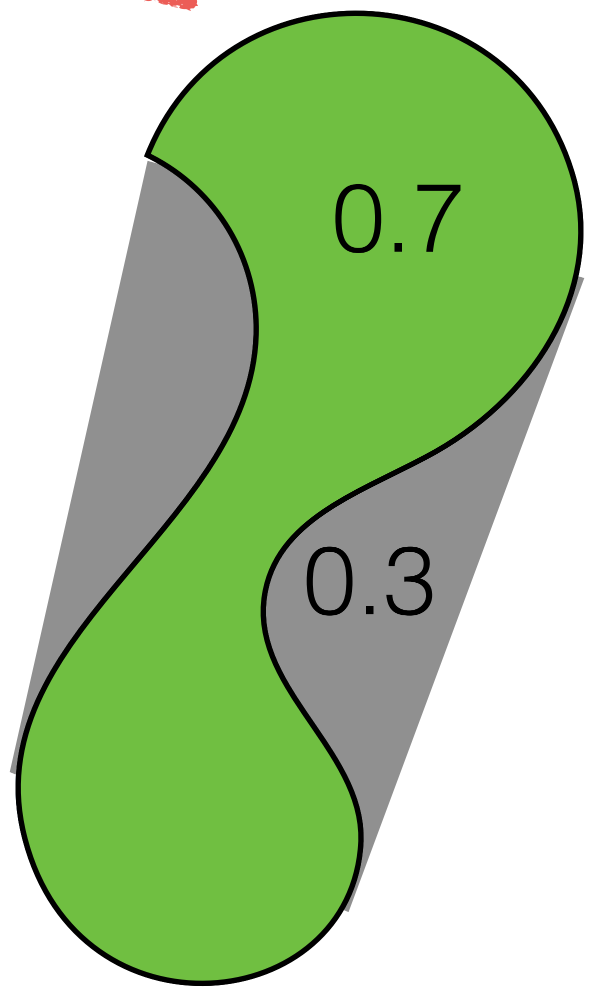
"Dataset"]

Out[202]=

	MaskedImage	Rectangularity	ConvexCoverage
1		0.496789	0.858025
2		0.539489	0.884615
3		0.522359	0.880503



why not 1?

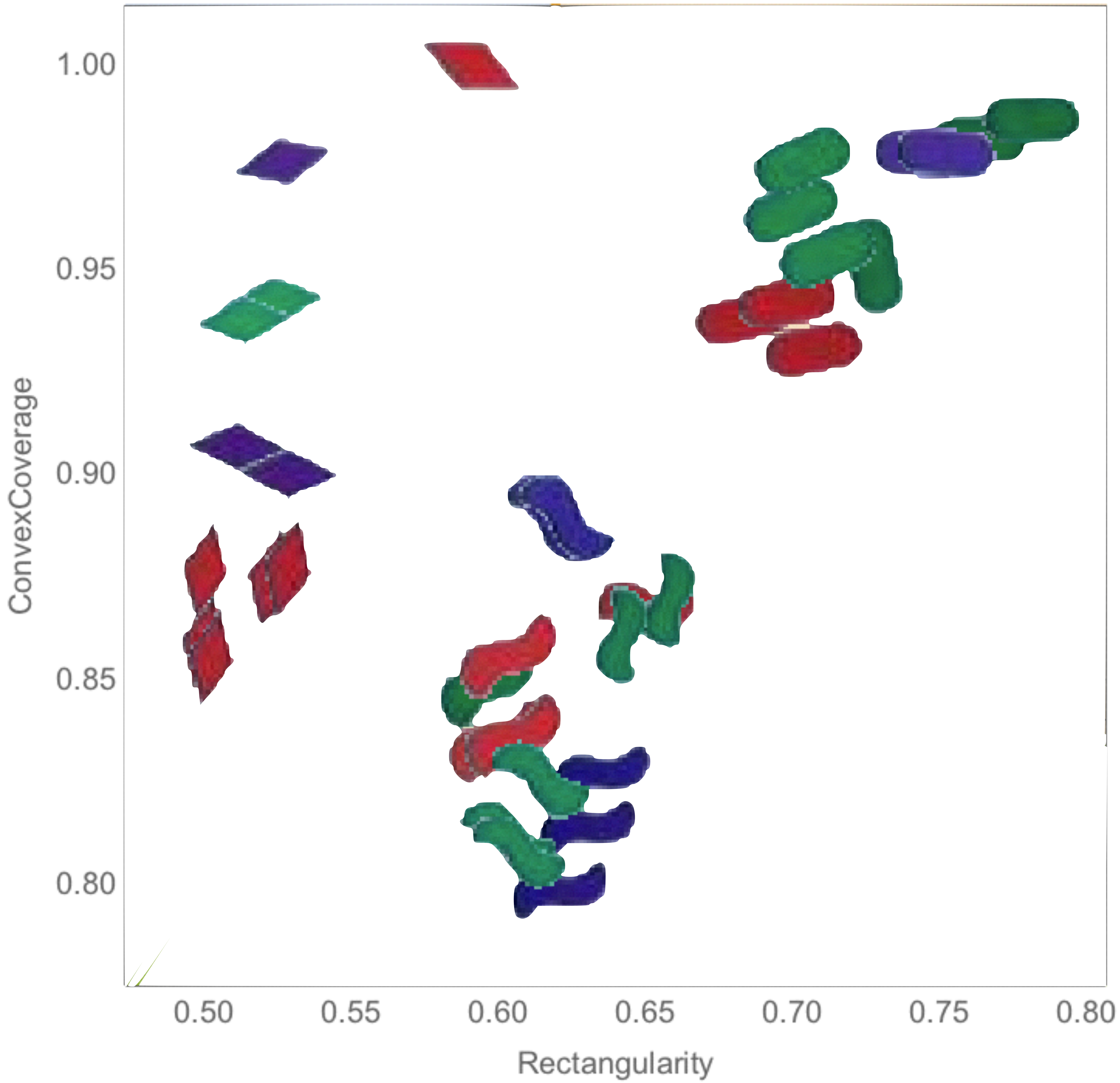


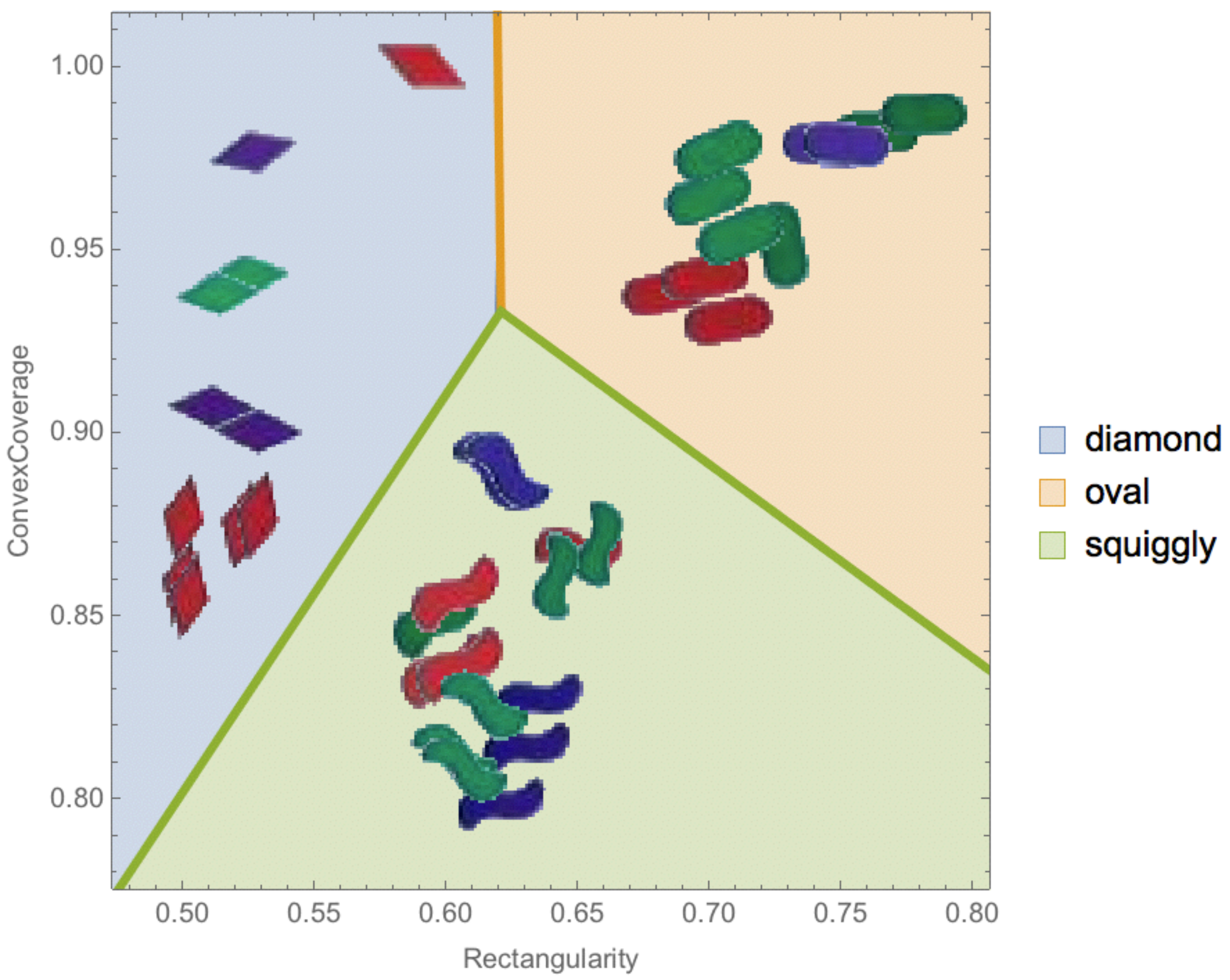
Out[202]=

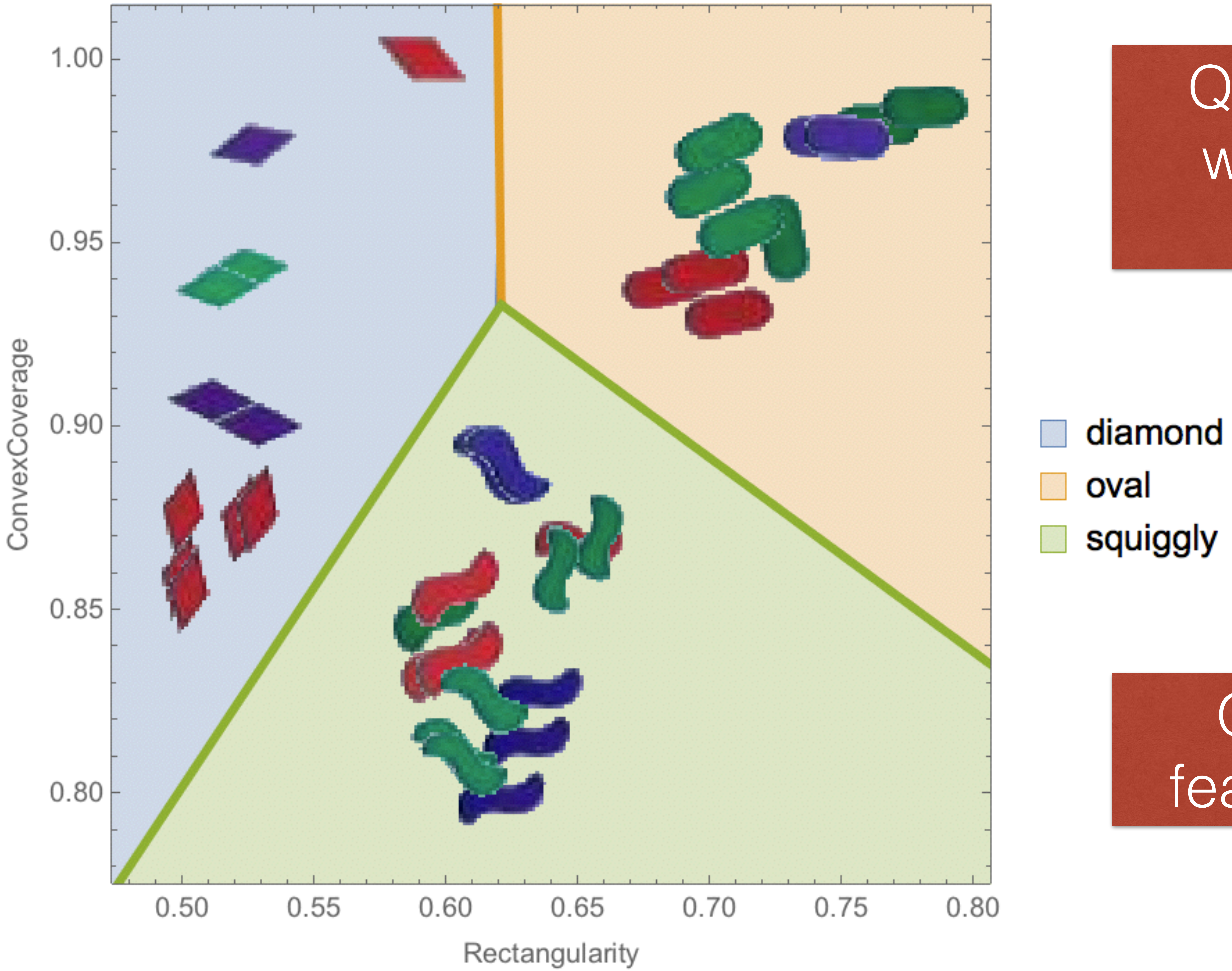
	MaskedImage	Rectangularity	ConvexCoverage
1		0.496789	0.858025
2		0.539489	0.884615
3		0.522359	0.880503

	MaskedImage	Rectangularity	ConvexCoverage
1		0.602219	0.838983
2		0.611343	0.844538
3		0.606565	0.860169

	MaskedImage	Rectangularity	ConvexCoverage
1		0.702335	0.968872
2		0.714635	0.954198
3		0.702327	0.977099


1.00
0.95
0.90
0.85
0.80
ConvexCoverage
0.50
0.55
0.60
0.65
0.70
0.75
0.80
Rectangularity

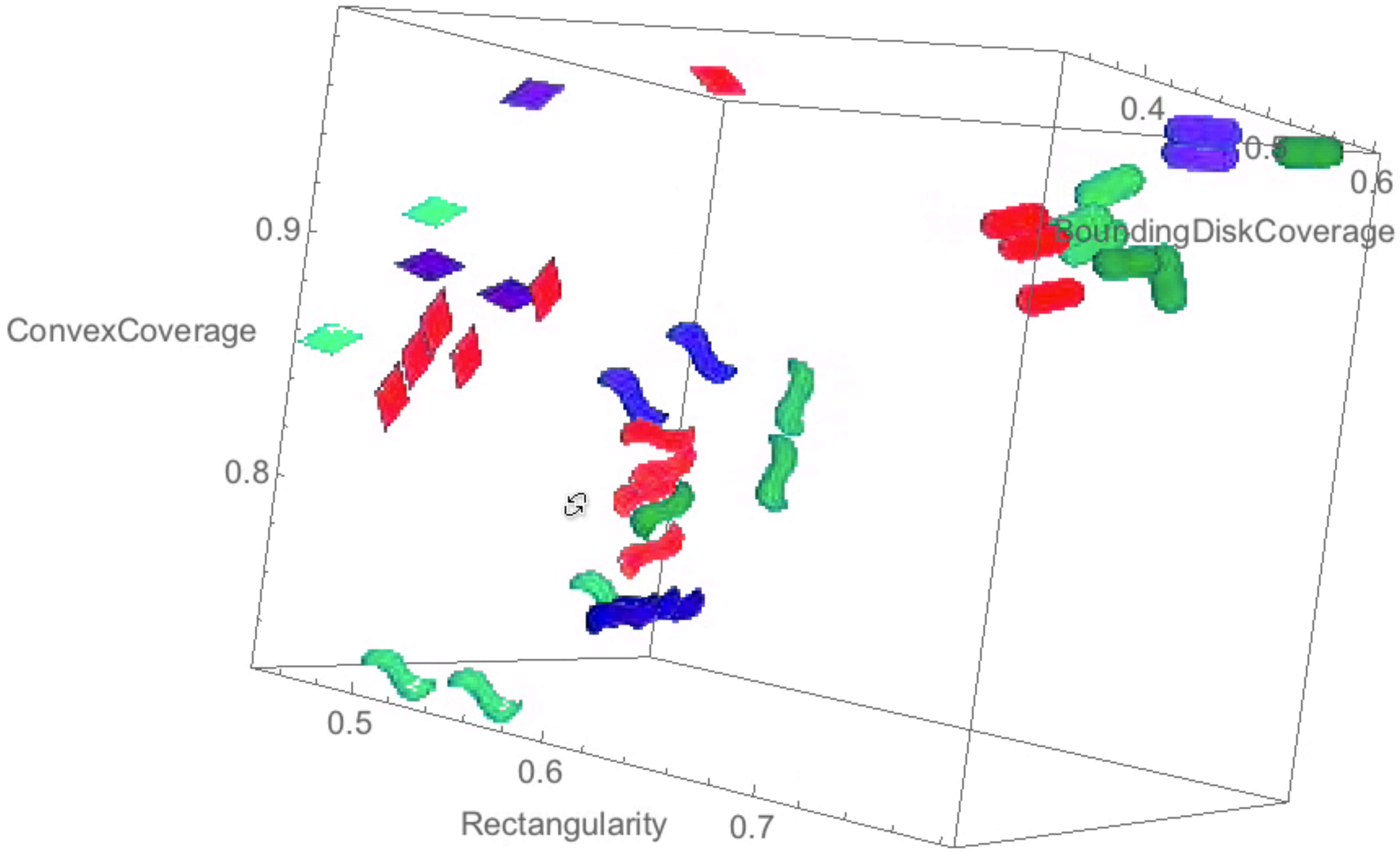

1.00
0.95
0.90
0.85
0.80
ConvexCoverage
0.50
0.55
0.60
0.65
0.70
0.75
0.80
Rectangularity
diamond
oval
squiggly



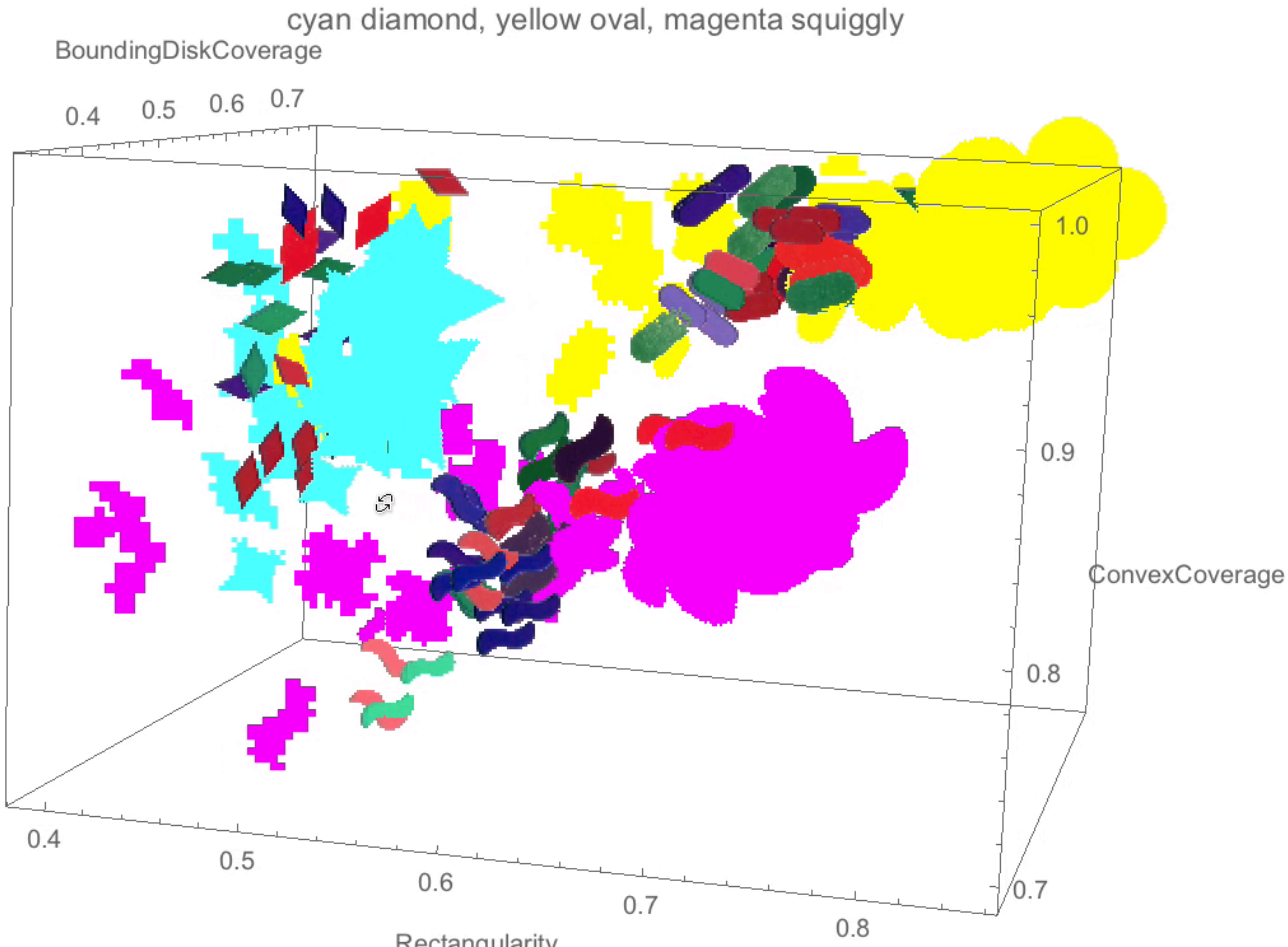
Q1: What if they weren't linearly separable?

Q2: Are these features "robust"?

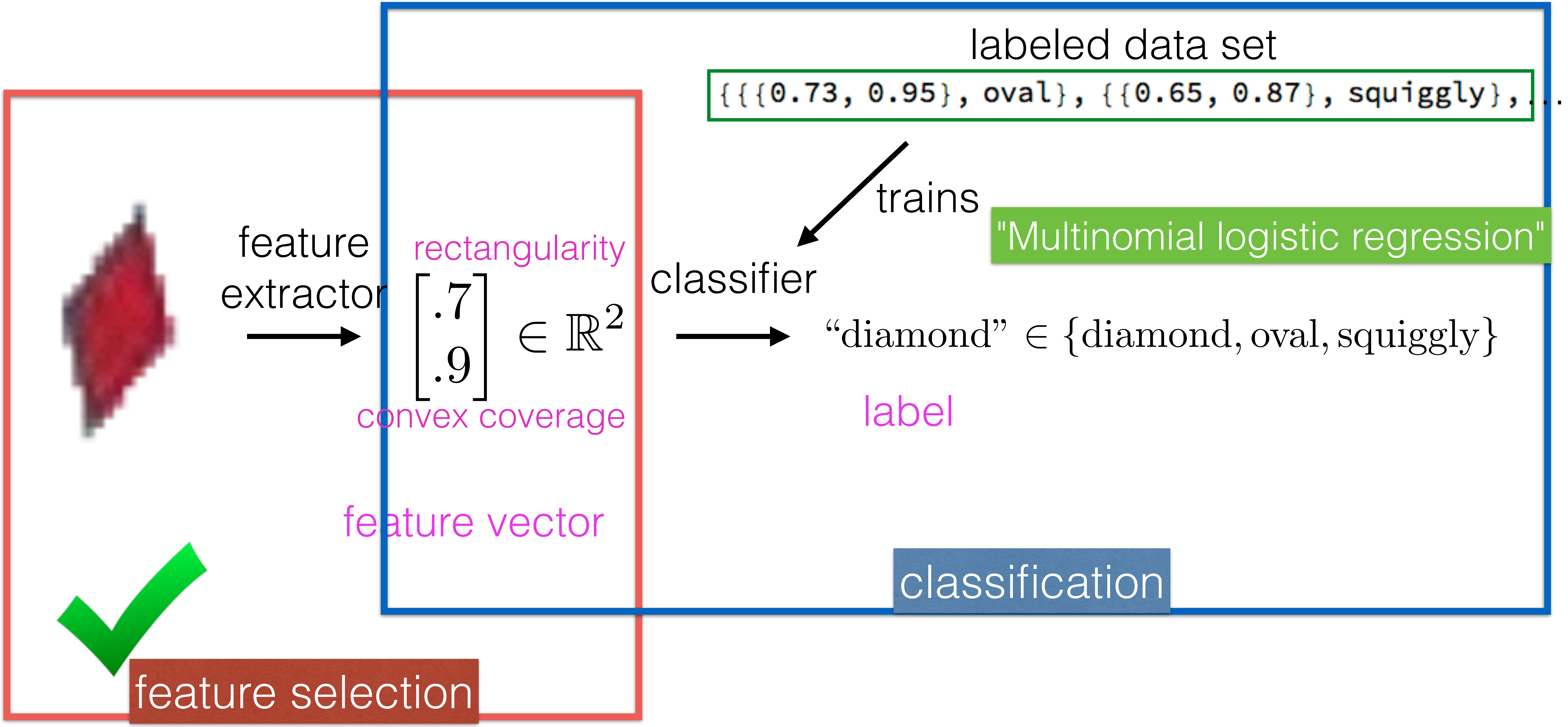
If not linearly separable, can add more dimensions ("lifting")

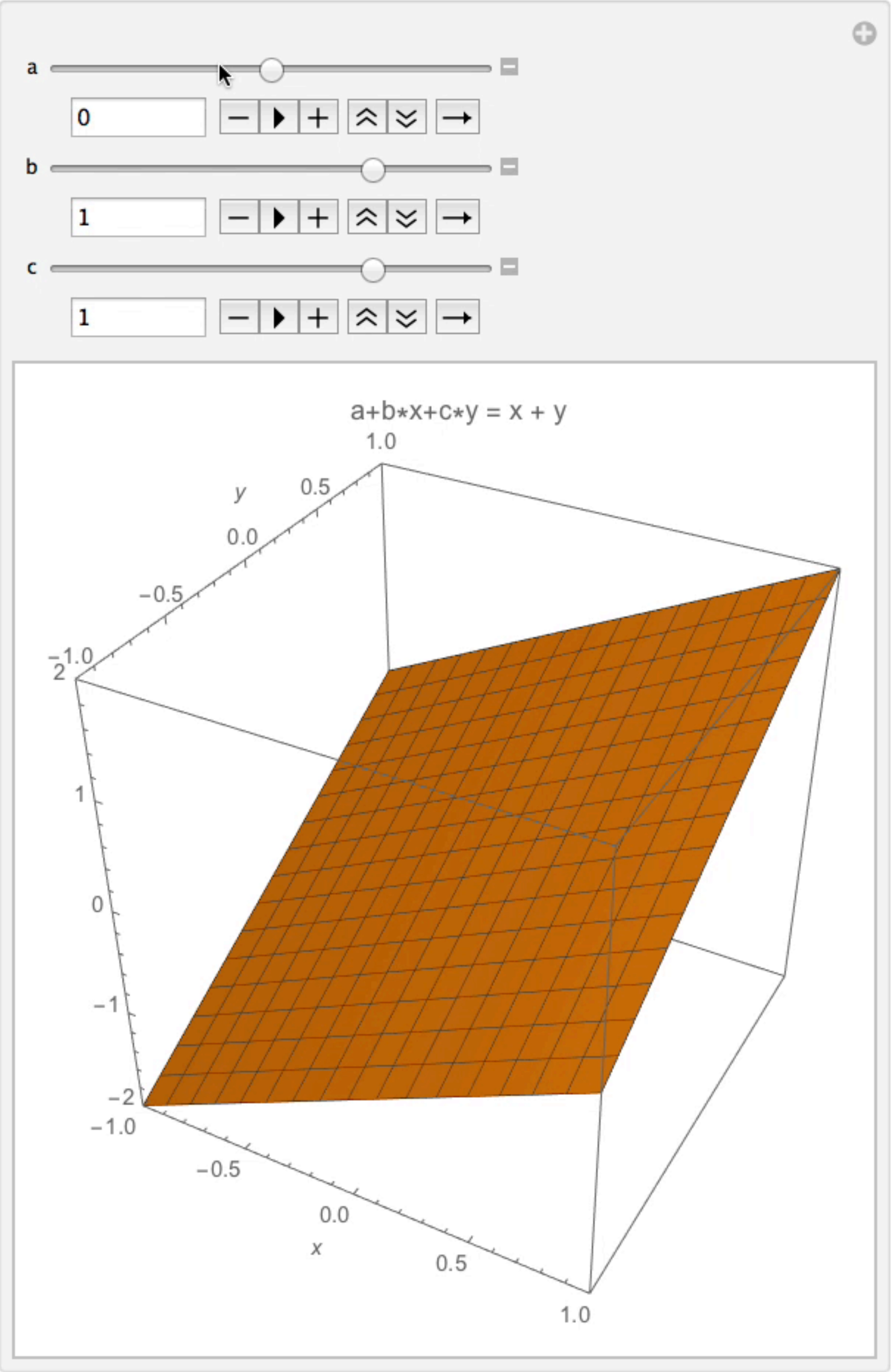


These 3 features are robust wrt rotation & scaling



Shape classification





Definition:

Affine function

$$f(x,y) = a + bx + cy$$

linear…

… plus a constant

$$f(x,y) = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} 1 \\ x \\ y \end{bmatrix}$$

Big Idea: affine function for each shape: its "probability"

Pick the shape with the greatest "probability"

Example: If measure

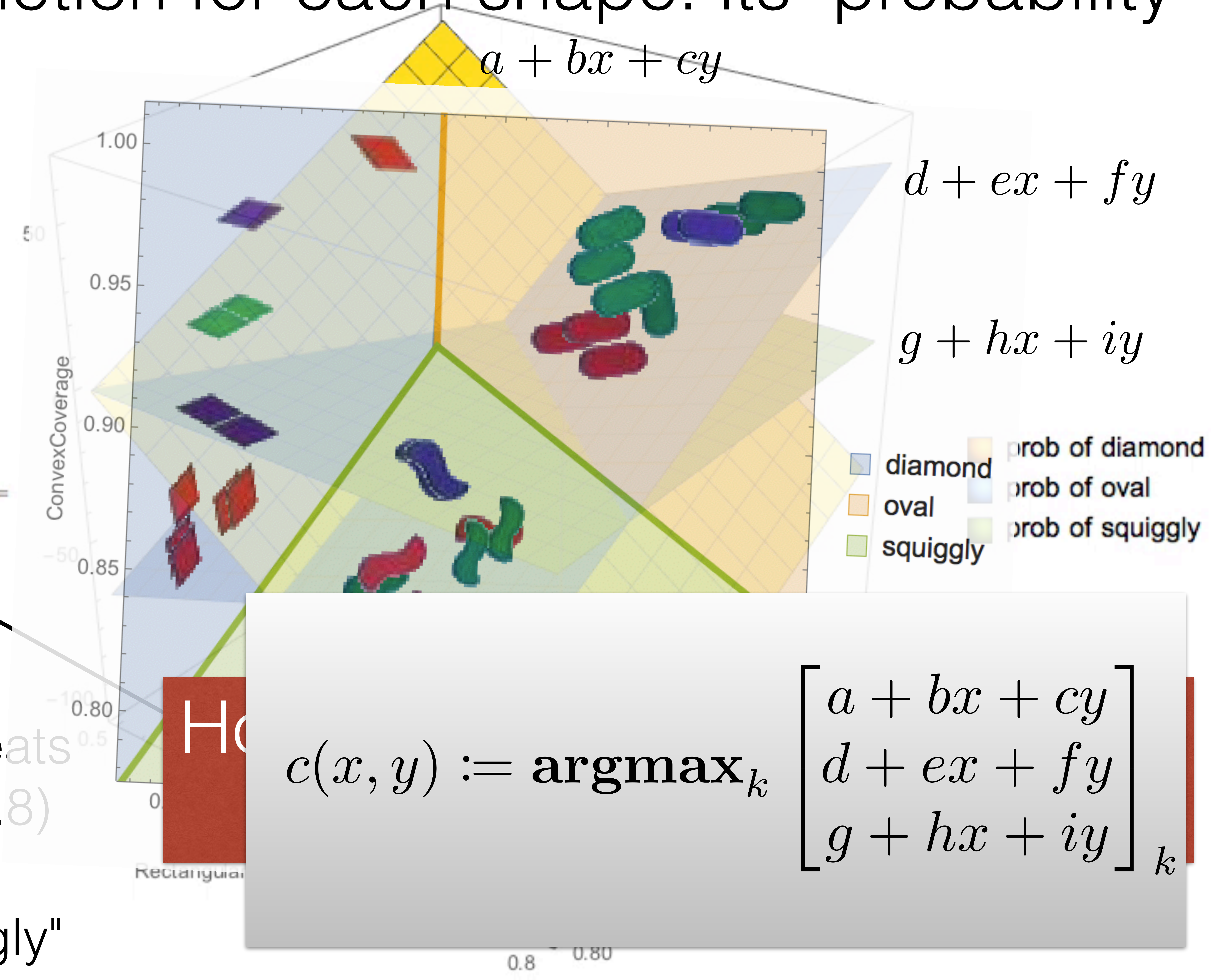
rectangularity=0.8
convex coverage=0.8

$\Downarrow$

green plane (squiggly) beats blue and yellow at (0.8,0.8)

$\Downarrow$

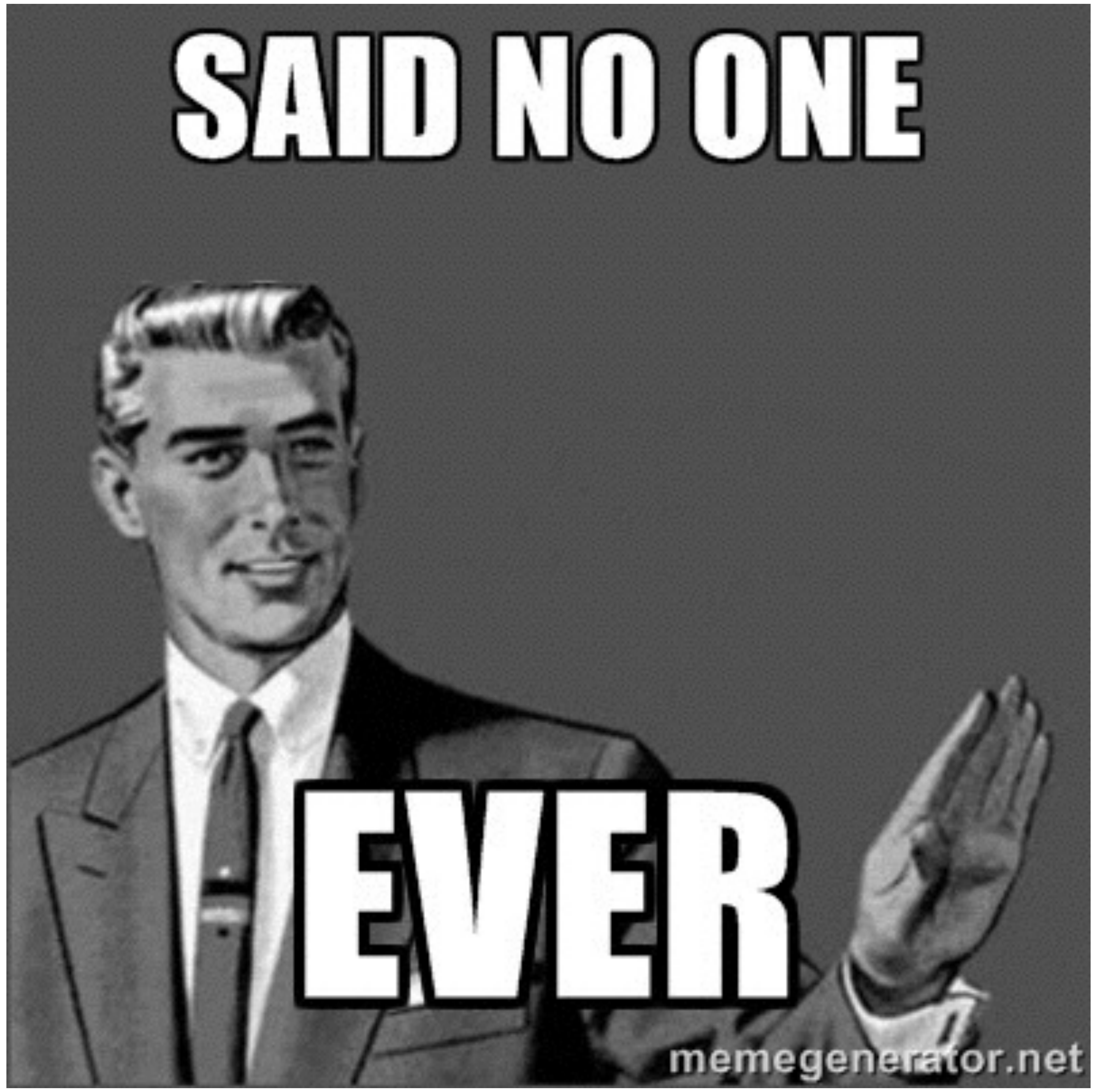
classify (0.8,0.8) as "squiggly"



$$c(x,y) := \mathbf{argmax}_k \begin{bmatrix} a + bx + cy \\ d + ex + fy \\ g + hx + iy \end{bmatrix}_k$$

Too hard to tune 9 variables.

Probability to the rescue



Assumption:

Given the feature vector *(x,y)*, the 3 shapes' probabilities are

rectangularity convex coverage

$$\mathbf{softmax}\left(\begin{bmatrix} a+bx+cy \\ d+ex+fy \\ g+hx+iy \end{bmatrix}\right)$$

(Make it a prob vector)

prob that (x,y) has label "diamond"

prob that (x,y) has label "oval"

prob that (x,y) has label "squiggly"

vec of same size

vec

$$\mathbf{softmax}(z) := \frac{e^z}{\sum e^z}$$

pos, sum to 1

Examples:

```
softmax[{-1, 2.5, -0.2}]
```

→ {0.012, 0.979, 0.009}

```
softmax[{1, 2, 3, 4, 5}]
```

→ {0.012, 0.032, 0.086, 0.234, 0.636}

Assumption:

Given the feature vector *(x,y)*, the 3 shapes' probabilities are

rectangularity convex coverage

$$\textbf{softmax}\left(\begin{bmatrix} a+bx+cy \\ d+ex+fy \\ g+hx+iy \end{bmatrix}\right)$$

(Make it a prob vector)

prob that (x,y) has label "diamond"

prob that (x,y) has label "oval"

prob that (x,y) has label "squiggly"

Example

For affine functions given by **a=1,b=2,…,i=9**,
the probability that the point (0.8,0.8) is "oval" is

```
In[570]:= With[{a = 1, b = 2, c = 3, d = 4, e = 5, f = 6, g = 7, h = 8, i = 9, x = .8, y = .8},
  softmax[{a + b x + c y, d + e x + f y, g + h x + i y}]
 ]
```

diamond oval squiggly

Out[570]= $\{1.67814 \times 10^{-7}, \boxed{0.000409567}, 0.99959\}$

Assumption:

Given the feature vector *(x,y)*, the 3 shapes' probabilities are

rectangularity convex coverage

$$\mathbf{softmax}\left(\begin{bmatrix} a+bx+cy \\ d+ex+fy \\ g+hx+iy \end{bmatrix}\right)$$

(Make it a prob vector)

prob that (x,y) has label "diamond"

prob that (x,y) has label "oval"

prob that (x,y) has label "squiggly"

Then the probability of observing the (labeled) data set

$$\{(x_1, y_1, lab_1), \ldots, (x_m, y_m, lab_m)\}$$

these are specifc numbers & labels

is

$$\prod_{k=1}^{m} \mathbf{softmax}\left(\begin{bmatrix} a+bx_k+cy_k \\ d+ex_k+fy_k \\ g+hx_k+iy_k \end{bmatrix}\right)_{lab_k}$$

eg: if lab_k = oval, then get 2nd elem

Assumption:

Given the feature vector *(x,y)*, the 3 shapes' probabilities are

rectangularity convex coverage

$$\textbf{softmax}\left(\begin{bmatrix} a + bx + cy \\ d + ex + fy \\ g + hx + iy \end{bmatrix}\right)$$

(Make it a prob vector)

prob that (x,y) has label "diamond"

prob that (x,y) has label "oval"

prob that (x,y) has label "squiggly"

Then the probability of observing the (labeled) data set

$$\{(x_1, y_1, lab_1), \ldots, (x_m, y_m, lab_m)\}$$

these are specifc numbers & labels

is $$L(a, b, \ldots, h, i) := \prod_{k=1}^{m} \textbf{softmax}\left(\begin{bmatrix} a + bx_k + cy_k \\ d + ex_k + fy_k \\ g + hx_k + iy_k \end{bmatrix}\right)_{lab_k}$$

"likelihood function"

eg: if lab_k = oval, then get 2nd elem

Then the probability of observing the (labeled) data set

$$\{(x_1, y_1, lab_1), \ldots, (x_m, y_m, lab_m)\}$$

these are specifc numbers & labels

is $$L(a, b, \ldots, h, i) := \prod_{k=1}^{m} \textbf{softmax}\left(\begin{bmatrix} a + bx_k + cy_k \\ d + ex_k + fy_k \\ g + hx_k + iy_k \end{bmatrix}\right)_{lab_k}$$

"likelihood function"

eg: if lab_k = oval, then get 2nd elem

Optimization problem:

these are specifc numbers & labels

Given labeled data set $\{(x_1, y_1, lab_1), \ldots, (x_m, y_m, lab_m)\}$,

find values for $a, b, \ldots, h, i \in \mathbb{R}$

that maximizes log ^ $L(a, b, \ldots, h, i)$

"maximum-likelihood estimation"

Optimization problem:

these are specifc numbers & labels

Given labeled data set $\{(x_1, y_1, lab_1), \ldots, (x_m, y_m, lab_m)\}$

find values for $a, b, \ldots, h, i \in \mathbb{R}$

"maximum-likelihood estimation"

log

that maximizes ^ $L(a, b, \ldots, h, i)$

```
In[627]:= data

Out[627]= {{0.727927, 0.951128, 2}, {0.650553, 0.868085, 3},
  {0.707464, 0.930736, 2}, {0.683217, 0.938596, 2}, ...
```

In[628]:= $\texttt{loglikelihood} = \sum_{x}^{\texttt{data}} \texttt{Log}\left[\texttt{softmax}\left[\begin{pmatrix} a + b\, x[[1]] + c\;\; x[[2]] \\ d + e\, x[[1]] + f\;\; x[[2]] \\ g + h\, x[[1]] + i\;\; x[[2]] \end{pmatrix}\right][[x[[3]]]]\right];$

```
In[629]:= FindMaximum[loglikelihood, {a, b, c, d, e, f, g, h, i}]
```

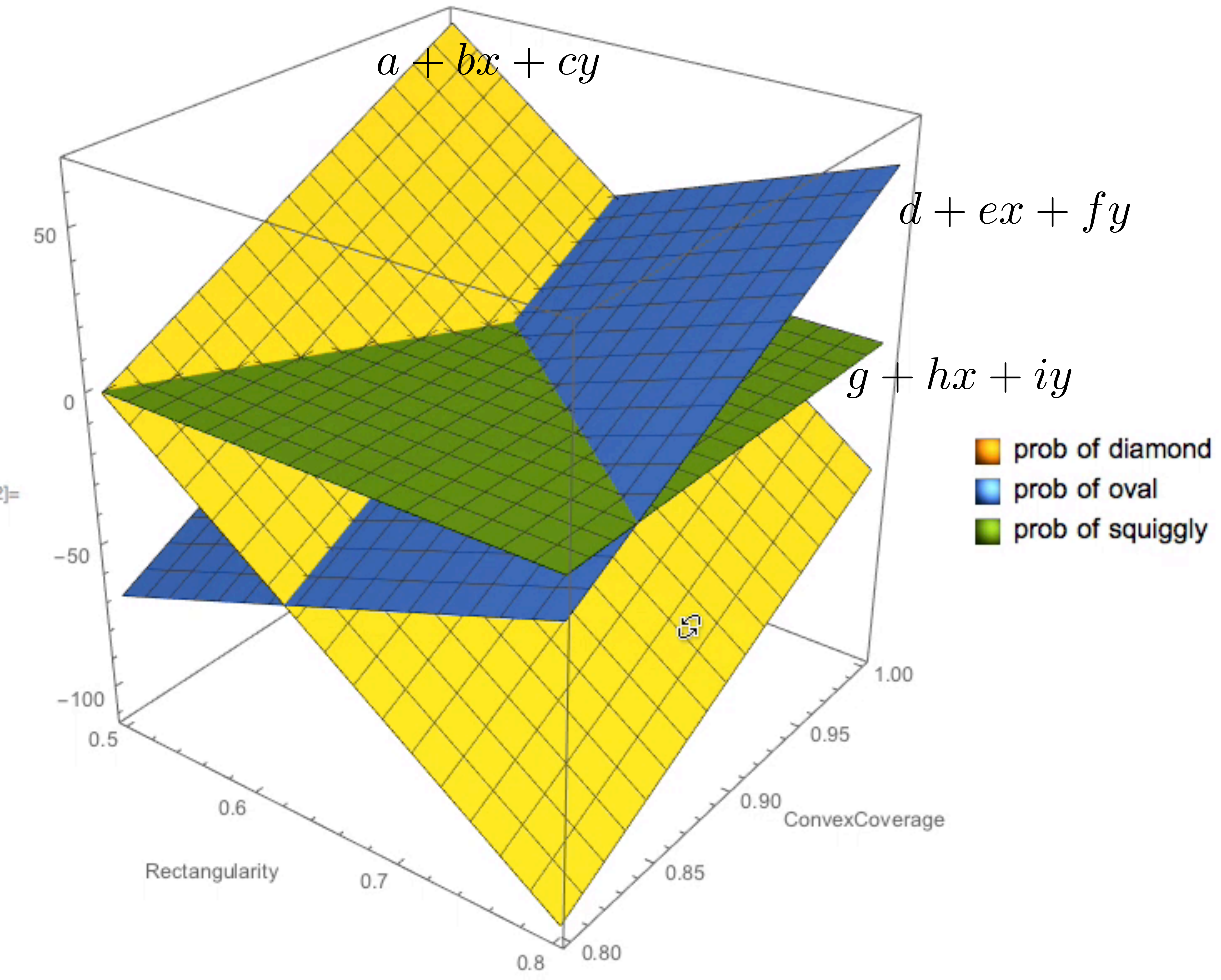
```
In[627]:= data

Out[627]= {{0.727927, 0.951128, 2}, {0.650553, 0.868085, 3}, ⋯
  {0.707464, 0.930736, 2}, {0.683217, 0.938596, 2},
```

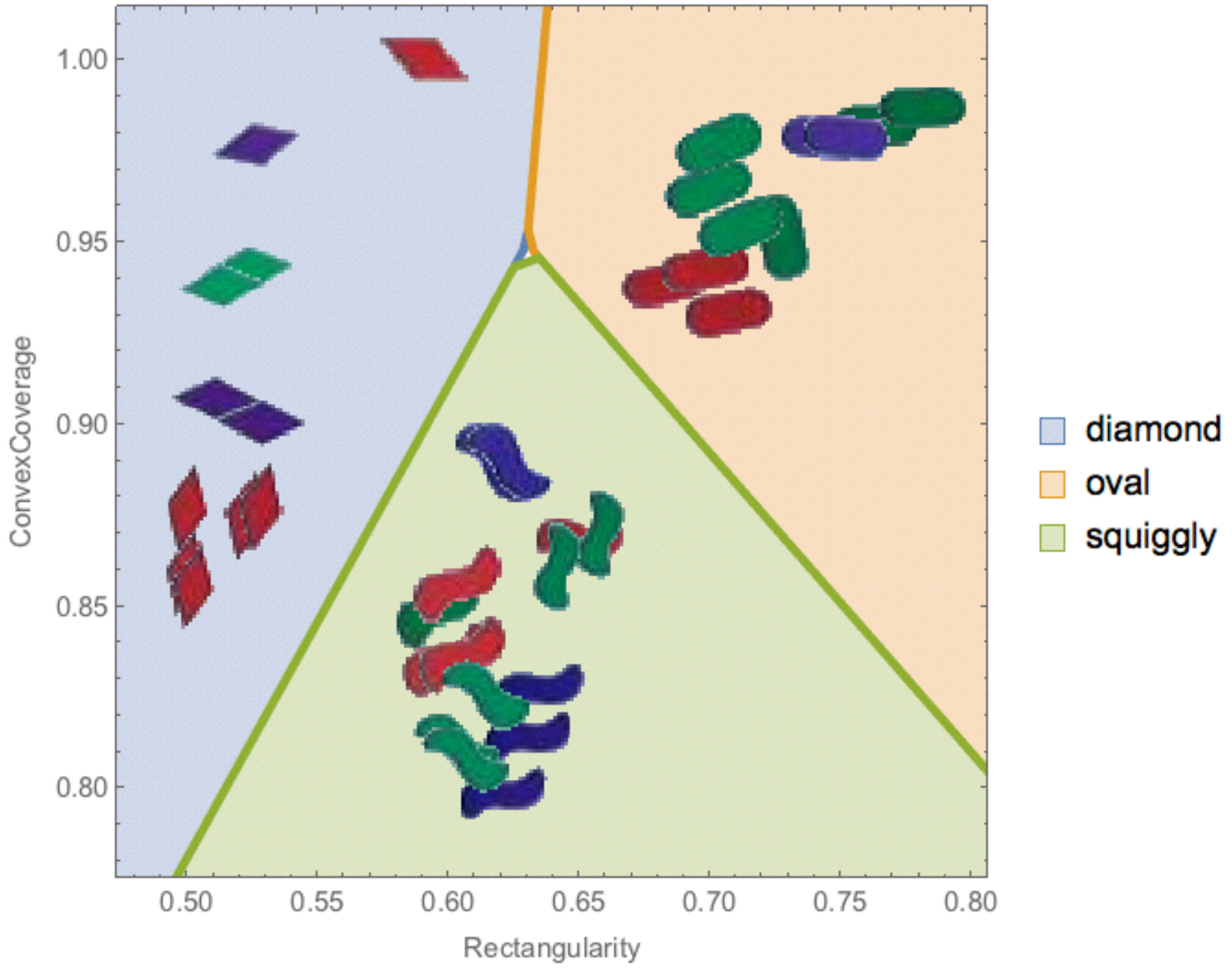
In[628]:= $\texttt{loglikelihood} = \sum_{x}^{\texttt{data}} \texttt{Log}\left[\texttt{softmax}\left[\begin{pmatrix} a + b\,x[[1]] + c\;\; x[[2]] \\ d + e\,x[[1]] + f\;\; x[[2]] \\ g + h\,x[[1]] + i\;\; x[[2]] \end{pmatrix}\right][[x[[3]]]]\right];$

```
In[629]:= FindMaximum[loglikelihood, {a, b, c, d, e, f, g, h, i}]
```

$\{a \to 4082.56,\ b \to -15\,232.7,\ c \to 5815.42,\ d \to -10\,019.5,\ e \to 12\,089.9,\ f \to 252$

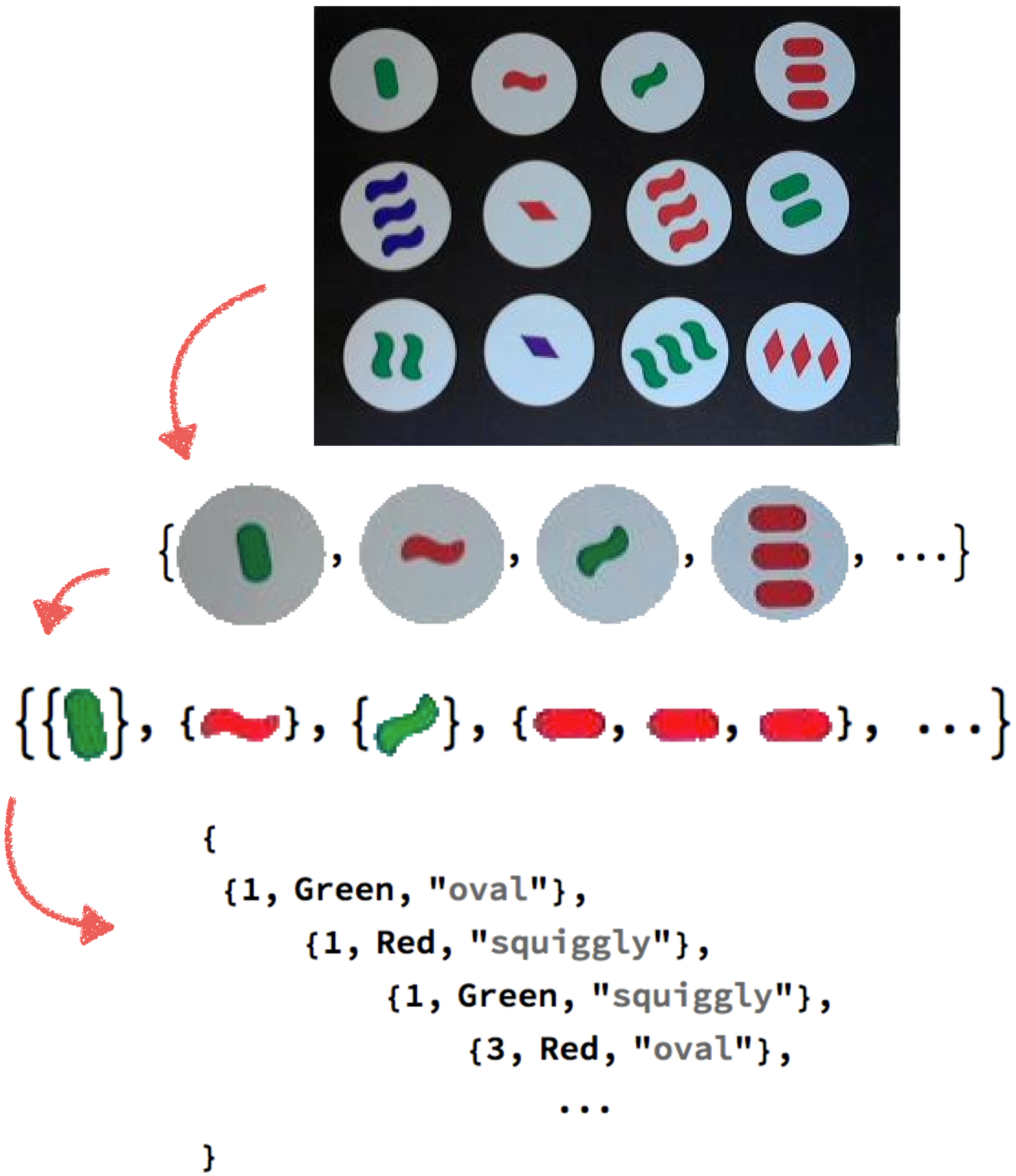
{a → 4082.56, b → -15 232.7, c → 5815.42, d → -10 019.5, e → 12 089.9, f → 252





Outline

- Set
- Mathematica
- Parsing the image ✓
  - Count ✓
  - Color ✓
  - Shape ✓
- Live demo **omg**



Wednesday, January 10 2018 12:54:54.976

playing...

Wed 10 Jan 2018 12:54:54 GMT-8.

no sets found


RESISTANCE
IS
FUTILE

END

Multinomial logistic regression

labeled data set

\# samples $m \approx 30$

\# features $n = 4$

\# classes $k = 3$

$$\text{Given } (x_i, y_i), \quad i = 1, \ldots, m \qquad x_i \in \mathbb{R}^n \qquad y_i \in \{\text{diamond}, \text{oval}, \text{squiggly}\} \ldots$$

measured feature-vector

label

$$\text{find } \theta_j \in \mathbb{R}^n, \quad j = 1, \ldots, k$$

objective variables

\# classes $k = 3$

$$\Theta := \begin{bmatrix} \longleftarrow & \theta_1^T & \longrightarrow \\ \longleftarrow & \theta_2^T & \longrightarrow \\ & \vdots & \\ \longleftarrow & \theta_k^T & \longrightarrow \end{bmatrix}$$

$k = 3$

$n = 4$

to minimize

cost function

$$l(\Theta) := \log \prod_{i=1}^{m} p(y_i | x_i; \Theta)$$

"log likelihood function"

prob of label y_i given that we measured feature vec x_i

$$p(y_j | x; \Theta) := \mathbf{softmax}(\Theta x)_j$$

vec

vec of same size

$$\mathbf{softmax}(z) := \frac{e^z}{\sum e^z}$$

pos, sum to 1

Example:

$$p(\text{oval} \,|\, \begin{bmatrix} .7 \\ .6 \end{bmatrix}; \Theta) := \mathbf{softmax}(\Theta * \begin{bmatrix} .7 \\ .6 \end{bmatrix})_2$$

y_i

x_i

oval is 2nd label

warning: I'm missing some constraints like th_k = 0