DEPARTAMENTO DE COMPUTACION

Facultad de Ciencias Exactas y Naturales - UBA

Trabajo Práctico 1

22 de mayo de 2017 Reconocimiento de Patrones

Integrante	LU	Correo electrónico
Christian Cuneo	755/13	chriscuneo93@gmail.com
Julián Bayardo	850/13	julian@bayardo.com.ar

Facultad de Ciencias Exactas y Naturales

Universidad de Buenos Aires

Ciudad Universitaria - Pabellón I

Intendente Güiraldes 2160 - C1428EGA

Ciudad Autónoma de Buenos Aires - Argentina

Tel/Fax: (54 11) 4576-3359

http://exactas.uba.ar

Índice

1. Ejercicio 1

En este ejercicio se pidió simplemente generar dos distribuciones gaussianas bi-variadas (quiere decir en dos dimensiones) y luego graficar las densidades de las mismas. Por ultimo se pide utilizar estas densidades como método de clasificación, separando el espacio en dos regiones, una para cada distribución.

Los parámetros utilizados para generar las distribuciones fueron los siguientes:

$$\mu_1 = (-2{,}5, 2{,}5)\mu_2 = (2, 2)$$

$$\sigma_1 = \begin{bmatrix} 9 & -7{,}2 \\ -7{,}2 & 9 \end{bmatrix} \sigma_2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

A continuación veremos los resultados de la experimentación:

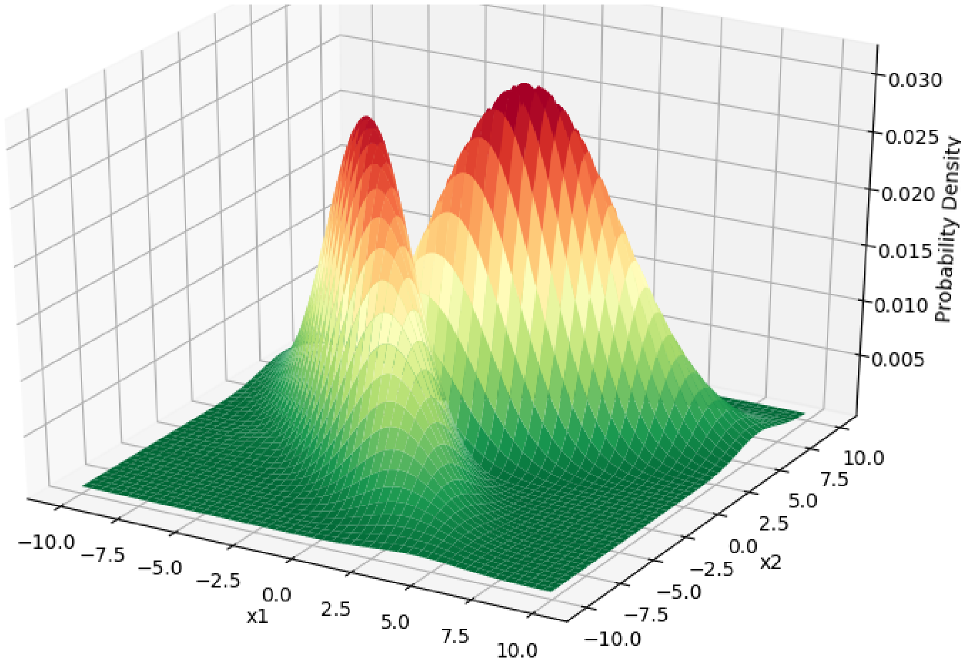

Figura 1: Densidades resultantes

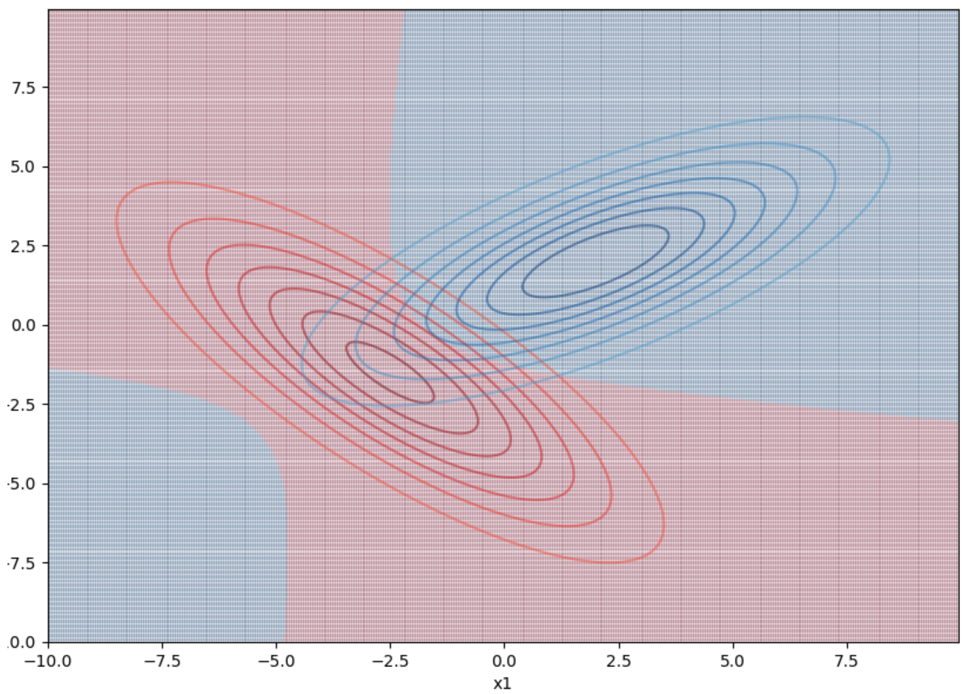

Figura 2: Areas resultantes de clasificacion gausseana

2. Ejercicio 2

Aquí se pidió generar imágenes sintéticas a partir de una imagen "phantom", quiere decir una imagen que muestra el resultado de una clasificación. Para esto vamos a necesitar generar datos aleatorios a partir de distribuciones gausseanas tri-variadas (por cada color RGB). Vamos a crear una distribución por cada clase, con media justamente en el valor medio de cada clase. A su ves necesitamos definir matrices de covarianza para cada una que describa la distribución de estos datos aleatorios. Para esto se nos pidió generar estrictamente tres conjuntos de matrices de covarianza y generar tres imágenes sintéticas a partir de ellas.

Por ultimo se pidió clasificar las imágenes generadas (luego de aplicarles un blur) y calcular la matriz de confusión

Primero se pidió utilizar una misma matriz de covarianza isotropica para todas las clases:

$$\sigma = \begin{bmatrix} 100 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 100 \end{bmatrix}$$

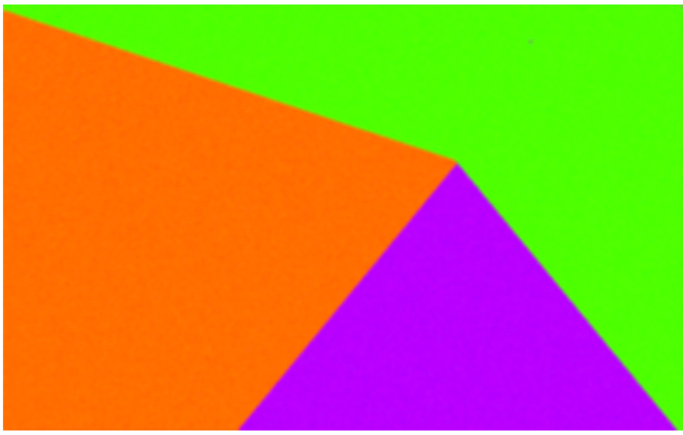

Figura 3: Imagen sintética generada caso I

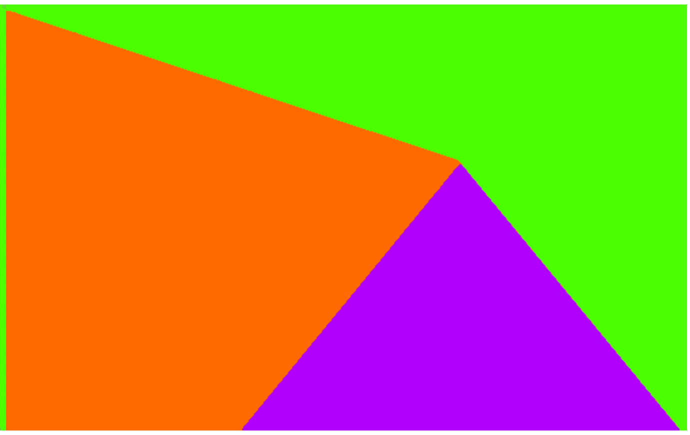

Figura 4: Resultado de clasificación caso I

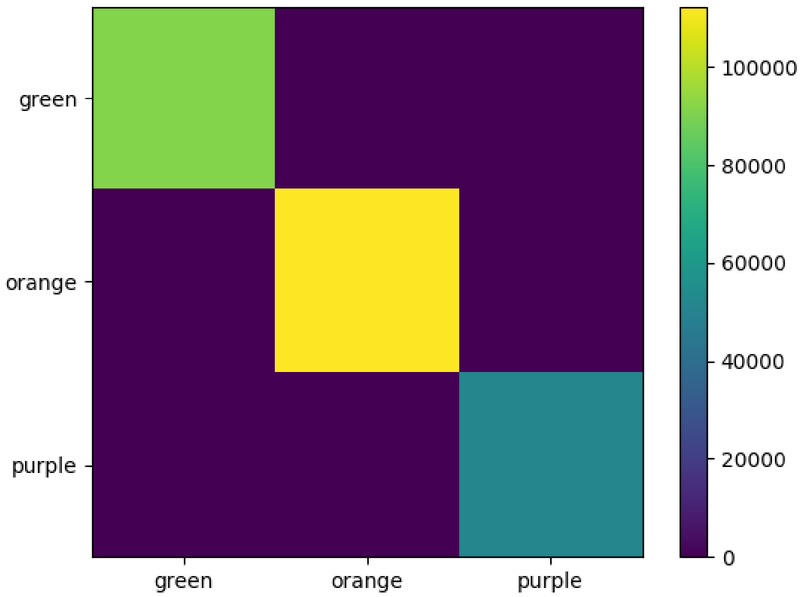

Figura 5: Matriz de covarianza caso I

Segundo se pidió utilizar una matriz diagonal diferente para cada clase:

$$\sigma_1 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 50 & 0 \\ 0 & 0 & 130 \end{bmatrix} \sigma_2 = \begin{bmatrix} 170 & 0 & 0 \\ 0 & 170 & 0 \\ 0 & 0 & 170 \end{bmatrix} \sigma_3 = \begin{bmatrix} 123 & 0 & 0 \\ 0 & 233 & 0 \\ 0 & 0 & 87 \end{bmatrix}$$

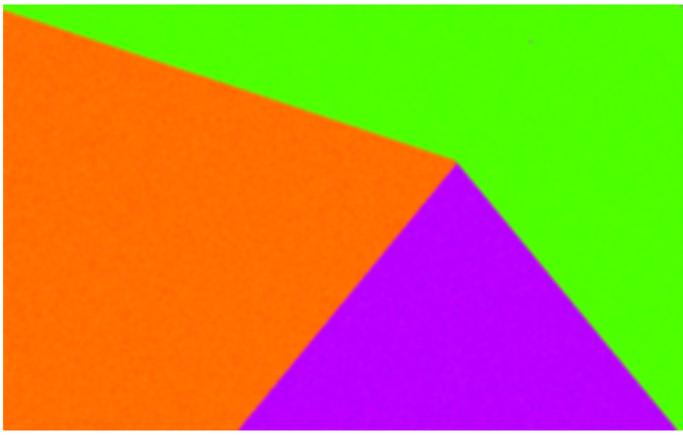

Figura 6: Imagen sintética generada caso II

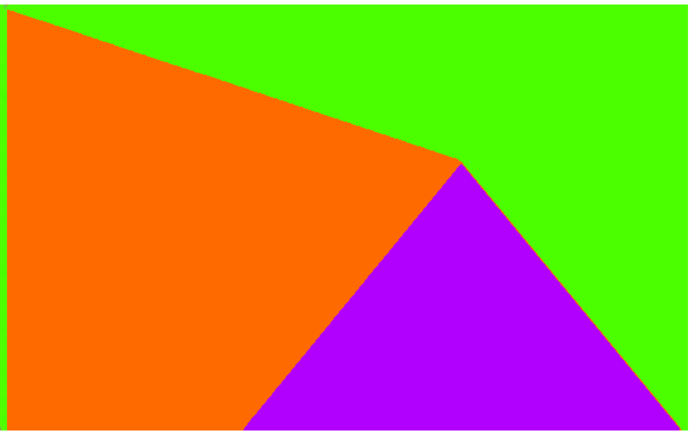

Figura 7: Resultado de clasificación caso II

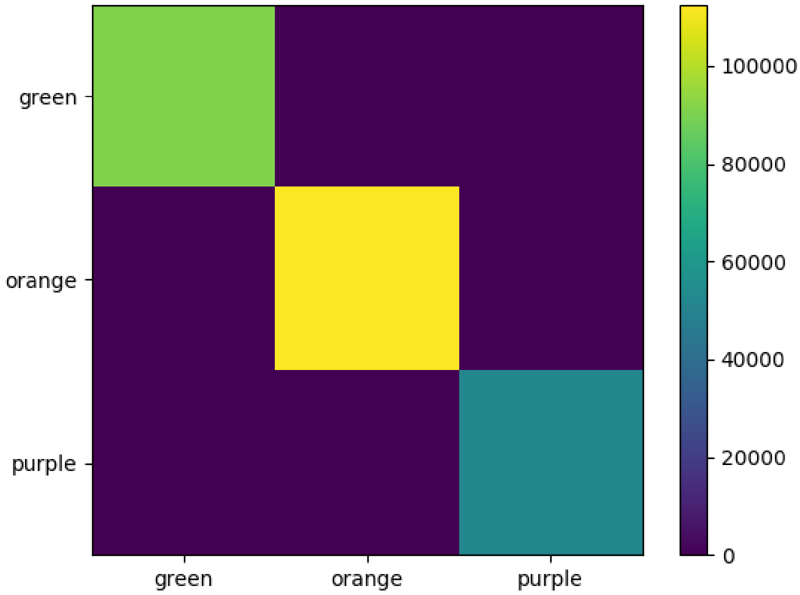

Figura 8: Matriz de covarianza caso II

Por tercero y ultimo se pidió utilizar una matriz no diagonal diferente para cada clase:

$$\sigma_1 = \begin{bmatrix} 100 & 10 & 1 \\ 10 & 100 & 10 \\ 1 & 10 & 100 \end{bmatrix} \sigma_2 = \begin{bmatrix} 211 & -11 & 0 \\ -14 & 221 & -132 \\ 0 & -112 & 221 \end{bmatrix} \sigma_3 = \begin{bmatrix} 1300 & -310 & 0 \\ -310 & 330 & -310 \\ 0 & -310 & 880 \end{bmatrix}$$

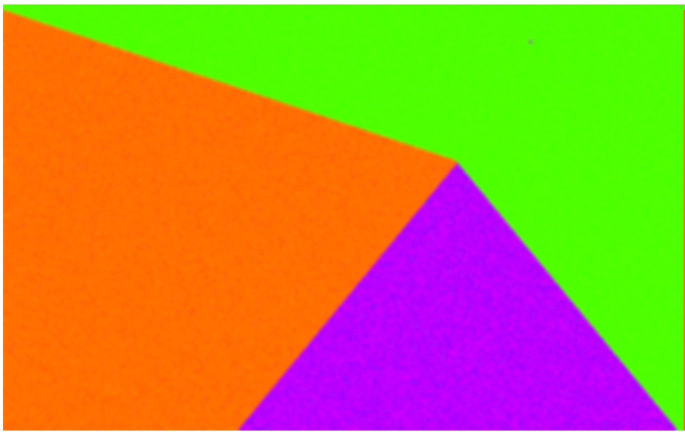

Figura 9: Imagen sintética generada caso III

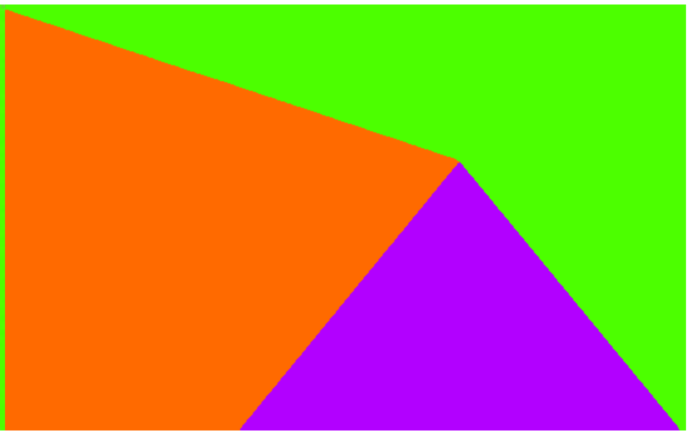

Figura 10: Resultado de clasificación caso III

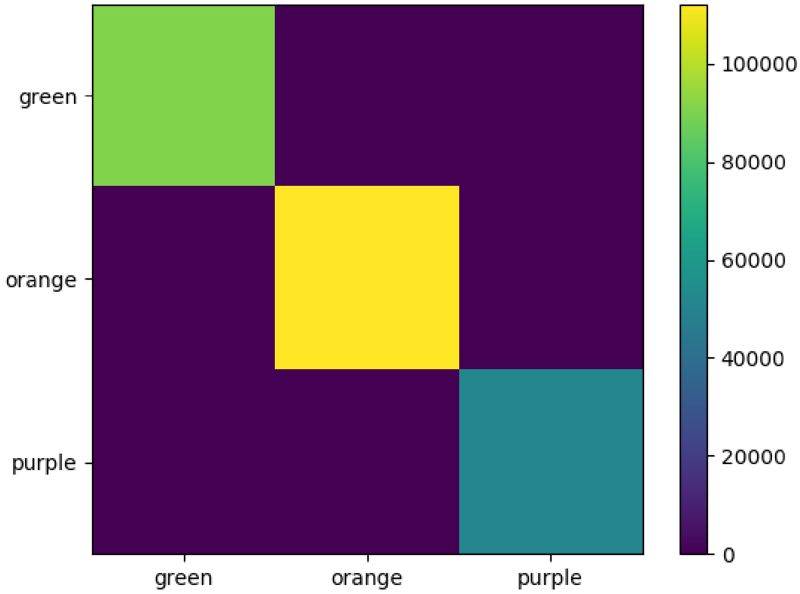

Figura 11: Matriz de covarianza caso III

Pudimos ver que las clasificaciones fueron extremadamente bien acertadas. Llegamos a la conclusión que, si a la hora de clasificar se usara otra matriz de covarianza para generar las funciones de densidades, la clasificación seria mas realista, ya que en la practica uno no suele saber como están distribuidos los datos desconocidos.

3. Ejercicio 3

En este ejercicio se nos presento una imagen satelital de unas siembras. La idea es primero reconocer a ojo las clases en lo que queremos clasificar a los pixeles de esta imagen. Luego etiquetar los pixeles de la imagen original en las clases definidas. Luego, a partir de esto datos, estimar los parámetros de una distribución gausseana tri-variada para cada clase. A continuación, utilizando estas distribuciones, clasificar la imagen original y ver cuan bien se logro clasificar la misma. Por ultimo, elegir zonas de prueba y ver cuan bien funcionaron las distribuciones generadas para clasificar estas áreas.

A ojo logramos distinguir tres clases: el desierto, siembra color rojo, y siembra (creemos seca) en color gris oscuro.

Veamos como etiquetamos estas clases:

Figura 12: Imagen original

Figura 13: Desierto

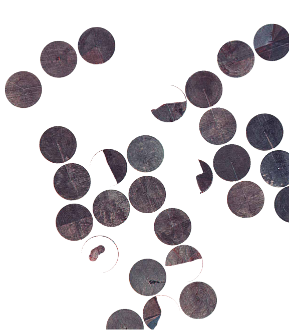

Figura 14: Siembra seca

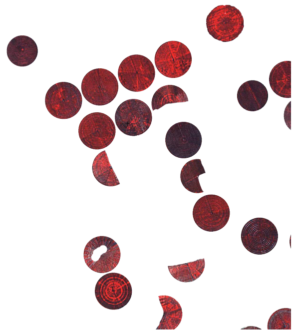

Figura 15: Siembra roja

Para estimar los parámetros de las distribuciones utilizamos solo un 80 de los datos de cada clase, el resto se va a utilizar al final como regiones de prueba. Estos fueron los resultados de la estimación:

$$\mu_1 = (158, 159, 155)\mu_2 = (115, 49, 57)\mu_3 = (92, 80, 90)$$

$$\sigma_1 = \begin{bmatrix} 1416 & 1280 & 1154 \\ 1280 & 1363 & 1223 \\ 1154 & 1223 & 1141 \end{bmatrix} \sigma_2 = \begin{bmatrix} 1757 & 427 & 352 \\ 427 & 373 & 358 \\ 352 & 358 & 393 \end{bmatrix} \sigma_3 = \begin{bmatrix} 672 & 556 & 494 \\ 556 & 563 & 506 \\ 494 & 506 & 477 \end{bmatrix}$$

Utilizando estos parámetros, clasificamos la imagen original utilizando las funciones de densidades para cada clase:

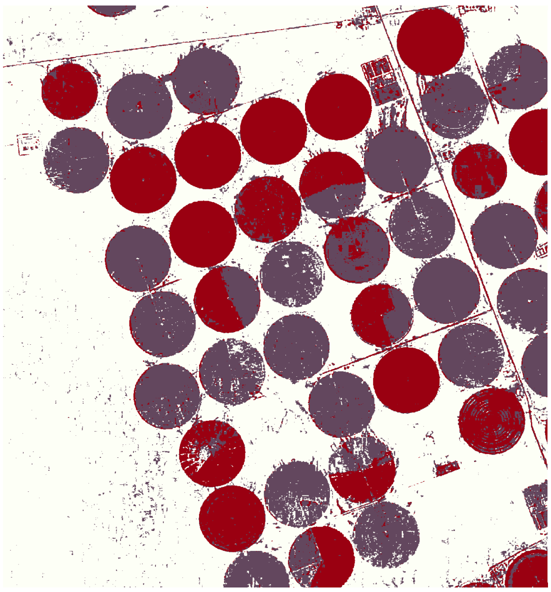

Figura 16: Resultado de clasificación

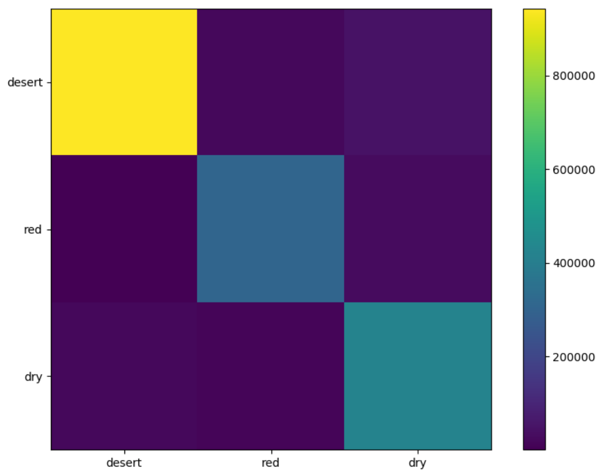

Figura 17: Matriz de confusión

Por ultimo clasificamos las zonas de prueba, que no fueron utilizadas para estimar, y obtuvimos los siguientes resultados:

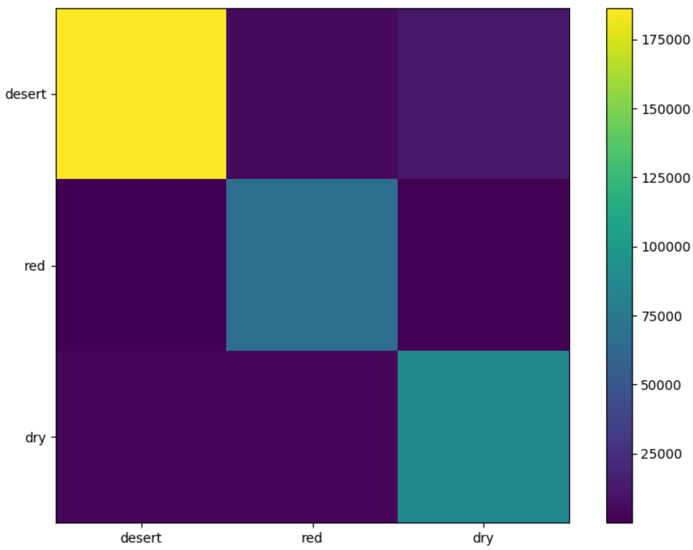

Figura 18: Matriz de confusión zonas de prueba