

Análisis y extracción de características de enfermedades de la piel: su aplicación en la detección de varicela

Virginia Arroyo y Julián Oyola

Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires

11 de abril de 2014

Contenido

1 Motivación

Contenido

1 Motivación

2 Las imágenes de piel y sus características

Contenido

- 1 Motivación
- 2 Las imágenes de piel y sus características
- 3 Detección de vesículas
 - Espacio de color
 - Detección de bordes
 - Detección de círculos
 - Ecualización del histograma

Contenido

- 1 Motivación
- 2 Las imágenes de piel y sus características
- 3 Detección de vesículas
 - Espacio de color
 - Detección de bordes
 - Detección de círculos
 - Ecualización del histograma
- 4 Discriminación entre varicela y otras enfermedades
 - Construcción de un histograma de referencia
 - Comparación entre una vesícula y la piel normal
 - Comparación entre vesículas de varicela y herpes zóster
 - Técnicas basadas en la comparación de media poblacional

Contenido

- 1 Motivación
- 2 Las imágenes de piel y sus características
- 3 Detección de vesículas
 - Espacio de color
 - Detección de bordes
 - Detección de círculos
 - Ecualización del histograma
- 4 Discriminación entre varicela y otras enfermedades
 - Construcción de un histograma de referencia
 - Comparación entre una vesícula y la piel normal
 - Comparación entre vesículas de varicela y herpes zóster
 - Técnicas basadas en la comparación de media poblacional
- 5 Conclusiones

¿Cómo detectar patrones de enfermedades de la piel?

- ¿Cómo detectar patrones de enfermedades de la piel?
- Resultan naturales y poco complejos para un ser humano
- ¿Y para un ordenador?

¿Cómo detectar patrones de enfermedades de la piel?

- ¿Cómo detectar patrones de enfermedades de la piel?
- Resultan naturales y poco complejos para un ser humano
- ¿Y para un ordenador?
- **Procesamiento digital de imágenes**

Ejemplo: Dos enfermedades



Varicela



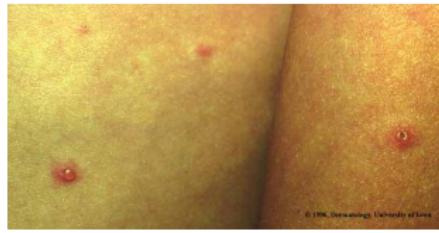
Herpes

Características de las imágenes utilizadas

Características de las imágenes utilizadas

- Heterogeneidad
- Escala
- Ruido
- Luminosidad
- Imperfecciones de la piel
- Elementos ajenos
- Distintas etapas de la enfermedad

Ejemplo: Variabilidad de imágenes para una misma enfermedad



Escala



© 1996, Dermatology, University of Iowa

Técnicas utilizadas y medidas adoptadas

Técnicas y medidas adoptadas

- Elección de un subconjunto de las imágenes

Técnicas utilizadas y medidas adoptadas

Técnicas y medidas adoptadas

- Elección de un subconjunto de las imágenes
- Ecualización del histograma (Contrast-limited adaptive histogram equalization)

Técnicas utilizadas y medidas adoptadas

Técnicas y medidas adoptadas

- Elección de un subconjunto de las imágenes
- Ecualización del histograma (Contrast-limited adaptive histogram equalization)
- Reducción del ruido o suavización utilizando un filtro gaussiano

Técnicas utilizadas y medidas adoptadas

Técnicas y medidas adoptadas

- Elección de un subconjunto de las imágenes
- Ecualización del histograma (Contrast-limited adaptive histogram equalization)
- Reducción del ruido o suavización utilizando un filtro gaussiano
- Elección del espacio de color

Algunas imágenes seleccionadas



© 1996, Dermatology, University of Iowa



© 1996, Dermatology, University of Iowa



© 1996, Dermatology, University of Iowa



© 1996, Dermatology, University of Iowa

Algunas imágenes seleccionadas

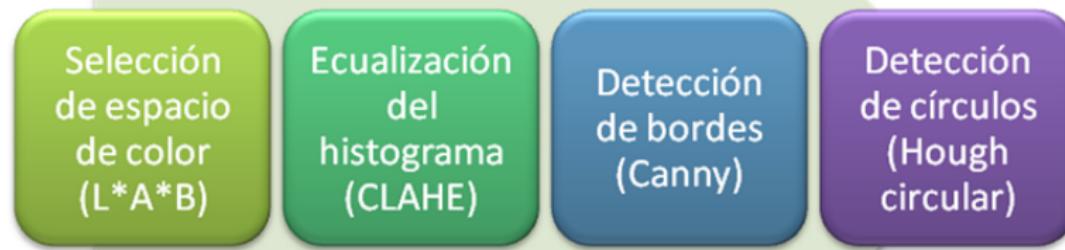


Metodología propuesta

Técnicas de procesamiento digital de imágenes

- Selección de espacio de color
- Detección de bordes
- Detección de círculos
- Ecualización del histograma

Esquema del proceso realizado



Espacio de color

¿Cómo representamos los colores y la luz en el ordenador?

- Espacios de color posibles
- Luminancia vs Crominancia
- YUV vs L^*a^*b

Luminancia: detección de bordes

Crominancia: detección de piel y falsos positivos

Luminancia vs Crominancia



© 1996, Dermatology, University of Iowa

Luminancia - componente L

Luminancia vs Crominancia



Crominancia - componente a

Luminancia vs Crominancia



Crominancia - componente b

Detección de bordes

- Resulta sencillo para el ser humano
- Borde: frontera entre el objeto y el fondo
- Existen varios métodos (Canny, Roberts, Sobel o Prewitt)
- Objetivos de un detector de borde:
 - Baja tasa de error
 - Buena localización del borde
- Método Canny:
 - Robusto contra el ruido
 - Gran adaptabilidad

Método de Canny

Criterios:

- Detectar los bordes sin eliminar ninguno y sin detectar bordes falsos
- Minimizar la diferencia entre la localización real del borde y la que detecta el algoritmo
- Evitar la detección de bordes duplicados

Método de Canny

Etapas del método:

- Suavizado de la imagen: Filtro gaussiano
- Obtención del gradiente: Filtro pasa altos en dirección vertical y horizontal
- Supresión de puntos que no son máximos locales:
Adelgazamiento del ancho de los bordes hasta lograr bordes de un píxel de ancho
- Umbral con histéresis: Función de histéresis basada en dos umbrales; con este proceso se trata de reducir la posibilidad de aparición de contornos falsos

Operaciones morfológicas

- Herramientas muy utilizadas en el procesamiento de imágenes
- Simplificar los datos de una imagen
- Preservar las características esenciales
- Eliminar aspectos irrelevantes
- Operaciones con la que probamos: open, close, clean, bridge

Ejemplo: Bordes detectados en algunas imágenes



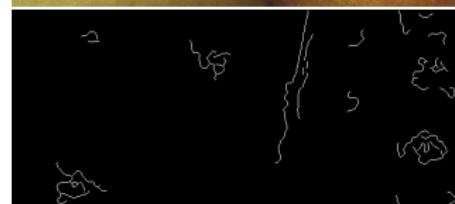
© 1996, Dermatology, University of Iowa



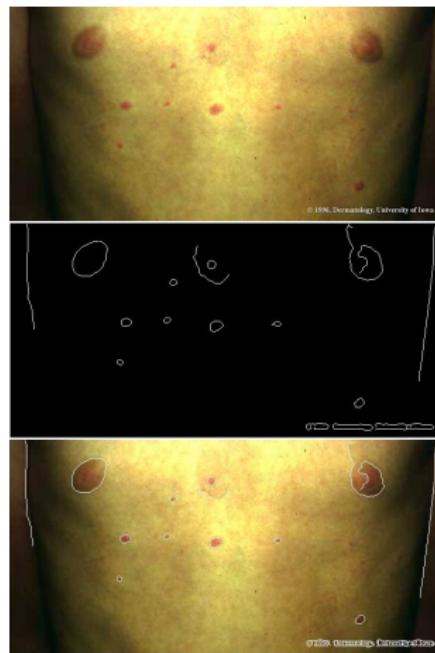
© 1996, Dermatology, University of Iowa



© 1996, Dermatology, University of Iowa



Ejemplo: Bordes detectados en algunas imágenes



Detección de círculos

- ¿Cómo determinar que estamos en presencia de un círculo?
- Transformada de Hough
 - Espacio de Hough
 - Arreglo de acumulación
- Definiciones de la transformada
 - Función de Parámetros de la Transformada
 - Transformada Directa de Hough
 - Transformada Inversa de Hough

Ejemplo: Arreglo de acumulación

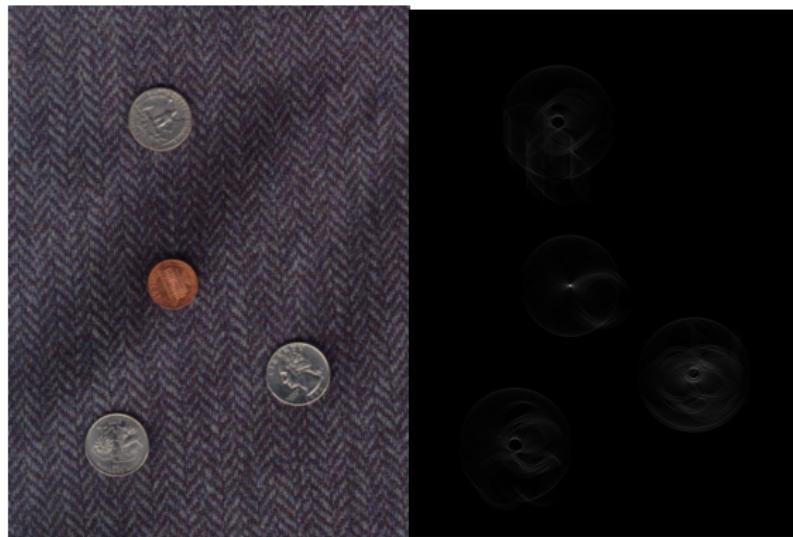


Figura: **Izquierda:** Fotografía original de las monedas. **Derecha:** Arreglo acumulador de votos.

Ejemplo: Arreglo de acumulación



Figura: **Izquierda:** Fotografía original de varicela. **Derecha:** Arreglo acumulador con votos sumados.

Ejemplo: Imagen con bordes detectados



© 1996, Dermatology, University of Iowa

Ejemplo: Imagen con el círculo detectado



© 1996, Dermatology, University of Iowa

Transformada Circular de Hough

Procedimiento realizado:

- Se aplica CHT: Circular Hough Transform. Ecuación
$$(x - x_0)^2 + (y - y_0)^2 = r^2$$
- Selección de candidatos: Ponderación con respecto al máximo y Umbralización
- Radios utilizados
- Eliminación de detecciones duplicadas

Radios Utilizados



Figura: Detección de vesículas con diferentes radios. **Izq:** Bordes detectados. **Centro:** Detección de vesículas con radio $\text{radio} = 40$. **Der:** Detección de vesículas con radio $\text{radio} = 43$.

Eliminación de detecciones duplicadas



Figura: Detección de vesículas con diferentes radios. **Izquierda:** radio = 23. **Derecha:** radio = 24.

Ecualización del histograma

¿Qué se desea obtener?

- Incrementar el contraste
- Mejorar la detección

CLAHE: Ecualización adaptativa limitando el contraste

Ejemplos de aplicación de ecualización

Sin aplicar ecualización



Figura: Detección de vesículas Izq: Bordes detectados. Der: Círculos detectados.

Ejemplos de aplicación de ecualización

Aplicando ecualización



Figura: Detección de vesículas Izq: Bordes detectados. Der: Círculos detectados.

Ejemplos de aplicación de ecualización

Sin aplicar ecualización



Figura: Detección de vesículas Izq: Bordes detectados. Der: Círculos detectados.

Ejemplos de aplicación de ecualización

Aplicando ecualización



Figura: Detección de vesículas Izq: Bordes detectados. Der: Círculos detectados.

Discriminación

¿Existen diferencias entre la piel sana y las vesículas?

¿Se puede detectar un falso positivo?

¿Existen diferencias entre las vesículas de la varicela y las de otra enfermedad?

¿Se pueden desarrollar técnicas de detección y discriminación automáticas?

- Análisis comparativo entre la varicela y la piel sana
- Análisis comparativo entre la varicela y otra enfermedad

Técnicas utilizadas

- Técnicas basadas en el histograma de color
 - Construcción de un histograma de referencia
 - Análisis de casos muestra: comparaciones entre el histograma de cada vesícula y
 - el histograma promedio de varicela
 - el histograma promedio de herpes zóster
- Técnicas basadas en la comparación de media poblacional
 - Comparación por distancia de Mahalanobis
 - Test de distribución Hotelling T-Cuadrada
 - Test de ANOVA

Histogramas sobre espacios de color

Análisis del histograma de las distintas componentes de color

- De forma individual
- Histograma bivariado de a pares de componentes

¿Qué espacio de color utilizar?

- L*a*b: Dos componentes de crominancia
- RGB: Información sobre el corrimiento al rojo de las vesículas en relación a la piel normal

Histogramas sobre espacios de color

Análisis del histograma de las distintas componentes de color

- De forma individual
- Histograma bivariado de a pares de componentes

¿Qué espacio de color utilizar?

- L*a*b: Dos componentes de crominancia
- RGB: Información sobre el corrimiento al rojo de las vesículas en relación a la piel normal

Histogramas de referencia: Valores extraídos de un conjunto de vesículas conocidas, contra los cuales comparar las nuevas vesículas

Histogramas

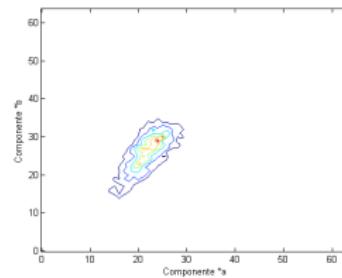
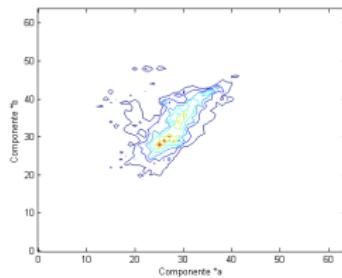


Figura: Histogramas bivariados promedio de varicela y herpes zóster, sobre las variables de crominancia $*a$ y $*b$. **Izq:** Varicela. **Der:** Herpes zóster.

Histogramas

- Leve corrimiento hacia el rojo del histograma de varicela
- Dispersión de valores mayor en el histograma del espacio RGB
- Mayor concentración de valores para el espacio de color L*a*b

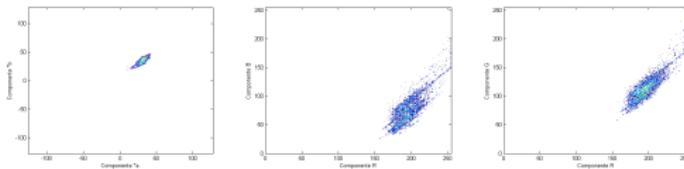


Figura: Histogramas bivariados promedio de la imagen izquierda de la figura, **Arriba izq:** , Variables de crominancia *a y *b. **Arriba der:** Variables de crominancia R y B. **Abajo:** Variables de crominancia R y B.

Divergencia entre histogramas

Mediciones de distancia (o divergencia) entre histogramas.

Indicador buscado: Si la distancia del histograma de la vesícula candidata al promedio está por debajo o por encima de cierto umbral.

Evaluación sobre:

- Norma 1
- Norma 2
- Divergencia de Kullback-Leibler (KLD)

Divergencia de Kullback-Leibler

Mide la diferencia entre dos distribuciones P y Q.

- P representa la distribución real de datos u observaciones (en este caso, la vesícula observada)
- Q representa el modelo teórico (en este caso, el histograma promedio)

Mide la cantidad esperada de bits extra requeridos para codificar ejemplos de P utilizando un código basado en Q.

También llamada “entropía relativa”.

Divergencia de Kullback-Leibler

$$KLD(P||Q) = \sum_i P(i) \ln\left(\frac{P(i)}{Q(i)}\right).$$

La KLD no es una distancia, porque no cumple con la condición de simetría.

Utilizamos una variante simétrica de la KLD:

$$KLD_{sym}(P||Q) = \frac{KLD(P||Q) + KLD(Q||P)}{2}.$$

Algunas imágenes analizadas

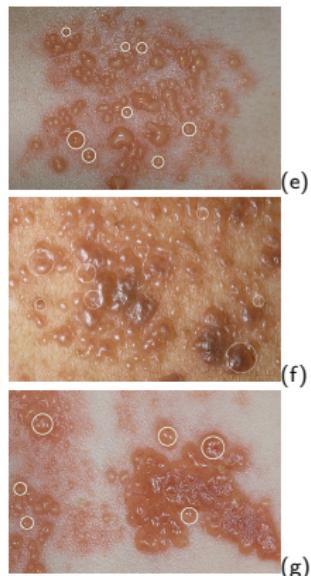
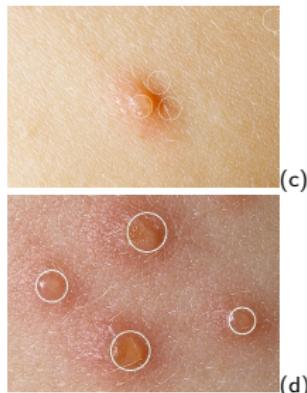
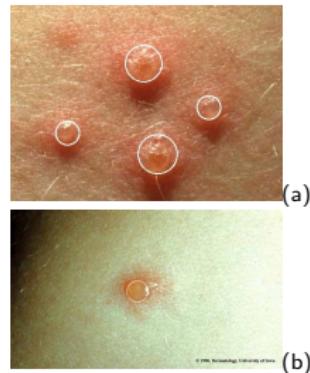


Figura: Vesículas analizadas. (a) - (b) - (c) - (d) Vesículas de varicela. (e) - (f) - (g) Vesículas de herpes zóster.

Resultados: Comparación entre varicela y piel

Los mejores resultados se observan sobre KLD_A , KLD_B y KLD_{AB} .

| Imagen | Círculo detectado | KLD_A | KLD_B | KLD_{AB} |
|------------|-------------------|---------|---------|------------|
| Varicela 1 | Falso positivo | 24,26 | 16,98 | 32,60 |
| Varicela 2 | Falso positivo | 27,54 | 17,22 | 45,41 |
| Varicela 3 | Falso positivo | 25,71 | 17,39 | 28,28 |
| Varicela 4 | Falso positivo | 22,52 | 15,74 | 24,90 |
| Varicela 1 | Vesícula 1 | 0,33 | 1,69 | 6,43 |
| Varicela 1 | Vesícula 2 | 0,51 | 1,16 | 7,15 |
| Varicela 1 | Vesícula 3 | 1,13 | 4,46 | 11,39 |
| Varicela 1 | Vesícula 4 | 0,68 | 2,60 | 11,17 |
| Varicela 2 | Vesícula 1 | 2,00 | 0,35 | 12,87 |
| Varicela 3 | Vesícula 1 | 2,17 | 3,87 | 12,66 |
| Varicela 3 | Vesícula 2 | 0,68 | 0,50 | 6,06 |
| Varicela 3 | Vesícula 3 | 1,49 | 0,49 | 7,01 |
| Varicela 3 | Vesícula 1 | 1,83 | 5,07 | 11,46 |

Cuadro: Mediciones de KLD entre histogramas de círculos detectados y los histogramas promedio, sobre componentes *a y *b.

Comparación entre vesículas de varicela y herpes zóster

Mediciones realizadas sobre:

- Un conjunto de círculos detectados de varicela
- Un conjunto de círculos detectados de herpes zóster
- Comparación contra el histograma promedio de la varicela

Resultado esperado:

- Divergencias de las vesículas de varicela: dentro un intervalo observable
- Divergencias de los histogramas de herpes: pocas o ninguna de las vesículas dentro de ese intervalo

Resultados: Comparación entre varicela y herpes zóster

Resultados no alentadores:

| Imagen 11 | Círculo detectado | KLD_A | KLD_B | KLD_{AB} |
|------------|-------------------|---------|---------|------------|
| Herpes 1 | Vesícula 1 | 7,86 | 10,42 | 18,73 |
| Herpes 1 | Vesícula 2 | 17,79 | 15,80 | 22,70 |
| Herpes 1 | Vesícula 3 | 7,24 | 11,72 | 20,27 |
| Herpes 1 | Vesícula 4 | 8,39 | 14,63 | 20,15 |
| Herpes 1 | Vesícula 5 | 7,42 | 10,35 | 17,85 |
| Herpes 1 | Vesícula 6 | 5,41 | 4,95 | 15,72 |
| Herpes 2 | Vesícula 1 | 4,03 | 5,78 | 12,86 |
| Herpes 2 | Vesícula 2 | 2,93 | 1,99 | 11,70 |
| Herpes 2 | Vesícula 3 | 6,35 | 6,79 | 15,60 |
| Herpes 3 | Vesícula 1 | 26,25 | 24,90 | 31,63 |
| Herpes 3 | Vesícula 2 | 4,81 | 9,57 | 19,64 |
| Varicela 1 | Vesícula 1 | 0,33 | 1,69 | 6,43 |
| Varicela 1 | Vesícula 2 | 0,51 | 1,16 | 7,15 |
| Varicela 1 | Vesícula 3 | 1,13 | 4,46 | 11,39 |
| Varicela 1 | Vesícula 4 | 0,68 | 2,60 | 11,17 |
| Varicela 2 | Vesícula 1 | 2,00 | 0,35 | 12,87 |
| Varicela 3 | Vesícula 1 | 2,17 | 3,87 | 12,66 |
| Varicela 3 | Vesícula 2 | 0,68 | 0,50 | 6,06 |
| Varicela 3 | Vesícula 3 | 1,49 | 0,49 | 7,01 |
| Varicela 3 | Vesícula 1 | 1,83 | 5,07 | 11,46 |

Cuadro: Mediciones de KLD entre círculos detectados de varicela y herpes zóster, contra el histograma promedio de varicela

Pruebas adicionales

- Medición de distancias: la norma 1 y norma 2
- Histogramas de componentes individuales de RGB
- Histogramas bivariados de componentes de RGB
- Histogramas de componentes individuales de L*a*b

En ninguna de estas pruebas se observó que el conjunto de mediciones de la varicela y del herpes zóster cayeran en intervalos disjuntos.

Conclusión parcial

- Para este conjunto de imágenes utilizadas, el análisis de la divergencia de histogramas entre la varicela y el herpes zóster no alcanza para determinar un umbral para distinguir entre ambas enfermedades.
- El mismo método arroja resultados favorables para determinar si una vesícula detectada es un falso positivo (piel) o una vesícula real de varicela.

Distancia de Mahalanobis

La distancia Mahalanobis entre dos vectores aleatorios multidimensionales x e y , provenientes de la misma población, se define como:

$$d_{mahal}(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}. \quad (1)$$

Siendo \vec{x} e \vec{y} los vectores entre los cuáles se desea medir la distancia.

Distancia de Mahalanobis

La distancia Mahalanobis entre un vector multivariado x a un conjunto de valores con media μ , y matriz de covarianza S , se define como:

$$D_{mahal}(x) = \sqrt{(x - \mu)^T S^{-1} (x - \mu)}. \quad (2)$$

$$x = (x_1, x_2, x_3, \dots, x_n)^T$$

$$\mu = (\mu_1, \mu_2, \mu_3, \dots, \mu_n)^T$$

Distancia de Mahalanobis

- Similar a la distancia euclídea
- Sirve para determinar la similitud entre los dos vectores
- A diferencia de la distancia euclídea, tiene en cuenta la correlación entre las variables
- Se adapta mejor para distribuciones simétricas no esféricas
- Es útil para determinar la distancia de un punto a la media de una distribución, o la distancia entre las medias de dos distribuciones
- Entonces, puede indicar si los píxeles de una vesícula detectada son más similares a los de varicela o a los de herpes zóster

Resultados: Comparación por distancia de Mahalanobis

Resultados favorables para discriminar los falsos positivos, pero no para distinguir entre vesículas de varicela y de herpes zóster.

| Imagen 11 | Círculo detectado | D_L | D_A | D_B | D_{LA} | D_{LB} | D_{AB} |
|------------|-------------------|-------|-------|-------|----------|----------|----------|
| Varicela 1 | Falso positivo 1 | 13,2 | 19,2 | 0,1 | 20,5 | 23,4 | 35,4 |
| Varicela 1 | Falso positivo 2 | 9,5 | 19,4 | 0,0 | 20,0 | 16,1 | 34,2 |
| Varicela 1 | Falso positivo 3 | 5,1 | 14,6 | 0,0 | 15,5 | 8,0 | 24,6 |
| Varicela 1 | Falso positivo 4 | 11,0 | 15,3 | 0,1 | 16,7 | 19,2 | 27,7 |
| Varicela 1 | Vesícula 1 | 1,5 | 3,0 | 0,9 | 5,3 | 3,0 | 3,7 |
| Varicela 1 | Vesícula 2 | 3,6 | 4,5 | 1,2 | 11,7 | 7,5 | 5,3 |
| Varicela 1 | Vesícula 3 | 1,7 | 1,6 | 0,3 | 3,5 | 2,4 | 2,0 |
| Varicela 1 | Vesícula 4 | 2,5 | 4,7 | 0,9 | 7,5 | 4,1 | 5,6 |
| Herpes 1 | Vesícula 1 | 0,90 | 2,70 | 1,60 | 6,29 | 4,19 | 3,70 |
| Herpes 1 | Vesícula 2 | 1,14 | 2,64 | 1,16 | 3,48 | 1,97 | 3,50 |
| Herpes 2 | Vesícula 1 | 2,94 | 2,69 | 1,00 | 13,53 | 8,38 | 3,15 |
| Herpes 2 | Vesícula 2 | 1,74 | 4,75 | 1,63 | 14,37 | 7,21 | 5,38 |
| Herpes 3 | Vesícula 1 | 1,39 | 6,83 | 1,95 | 17,51 | 6,48 | 7,53 |
| Herpes 3 | Vesícula 2 | 0,98 | 3,62 | 1,86 | 7,59 | 4,42 | 4,02 |

Cuadro: Mediciones de promedio de D_{mahal} entre círculos detectados de varicela, herpes zóster y falsos positivos, contra el histograma promedio de varicela

Distribución poblacional del color de la vesículas

- Distribuciones de color de ambas enfermedades muy similares
- ¿Pertenecen a una misma población?
- ¿Son muestras de la misma variable aleatoria?
- Pruebas estadísticas para determinar su homogeneidad
 - Test de distribución Hotelling T-Cuadrada
 - Test de ANOVA

Test de distribución Hotelling T-Cuadrada

- Asumiendo dos distribuciones multinormales y con matrices de covarianza idénticas, dada la hipótesis nula H_0 que ambas medias son iguales $\mu_1 = \mu_2$, puede indicar si hay suficiente evidencia para rechazar H_0 .
- Vector de diferencia utilizado: distancia de Mahalanobis.
 - Entre los píxeles de una vesícula de varicela y el histograma promedio de varicela
 - Entre los píxeles de una vesícula de herpes zóster y el histograma promedio de varicela

Resultados: Test de distribución Hotelling T-Cuadrada

- En todos los casos el test arrojó $p = 0$
- Indica que **no hay evidencia suficiente** para rechazar la hipótesis nula de que ambas distribuciones tienen la misma media.

Test de ANOVA

Evaluar si se puede rechazar la hipótesis nula H_0 de igualdad de las medias.

- Tomar imágenes representativas
- Construir una clase con los valores para las vesículas de cada imagen

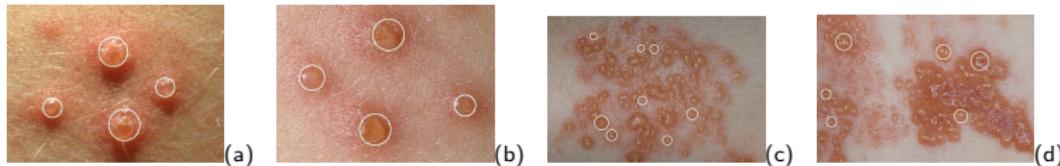


Figura: Vesículas detectadas. (a) - (b) Vesículas de varicela. (c) - (d) Vesículas de herpes zóster.

Resultados: Test de ANOVA

| Distancia entre medias | Intervalo | Comparación de medias |
|-----------------------------|--------------------|-----------------------|
| $\mu(a) - \mu(c) = 9,1967$ | [8,8831, 9,5104] | Interclases |
| $\mu(a) - \mu(d) = 6,1135$ | [5,8445, 6,3826] | Interclases |
| $\mu(b) - \mu(c) = 6,5746$ | [6,2689, 6,8802] | Interclases |
| $\mu(b) - \mu(d) = 3,4913$ | [3,2317, 3,7509] | Interclases |
| $\mu(a) - \mu(b) = 2,6222$ | [2,3835, 2,8609] | Intraclasses |
| $\mu(c) - \mu(d) = -3,0832$ | [-3,4131, -2,7533] | Intraclasses |

Cuadro: Comparación entre clases de varicela y herpes zóster, sobre componente *a

Trabajo futuro

- Detección de piel
- Detección de ampollas que no tengan forma circular
- Detección de patrones en las imágenes
 - Buscar características que permitan determinar cuándo se está en presencia de la varicela
 - Aprendizaje automático sobre el histograma del color de las ampollas detectadas

Preguntas?

...

Gracias!

Virginia Arroyo (virginia.arroyo@gmail.com)
Julián Oyola (joyola@dc.uba.ar)

Directores:

Anita Ruedin (ana.ruedin@dc.uba.ar)
Daniel Acevedo (dacevedo@dc.uba.ar)