

Clasificador de Páginas Web Pornográficas Basado en el Contenido de las Imágenes

William Armando Ceballos* Luis Eduardo Salazar* Ana
Isabel Oviedo†

Fecha de Recibido: 07/11/2008 Fecha de Aprobación: 20/04/2009

Resumen

La World Wide Web, o simplemente web, es un sistema lógico de acceso y búsqueda de información disponible en Internet cuyas unidades informativas son las páginas web. La web ha facilitado la publicación de gran cantidad de información accesible desde cualquier lugar del mundo; sin embargo, parte del contenido ofrecido como la pornografía, es considerado inapropiado para algunos usuarios.

Para aportar al filtrado de pornografía en la web, este trabajo propone el desarrollo de un clasificador de páginas web basado en la evaluación de las imágenes presentes en el contenido de la página. La evaluación de las imágenes es realizada en tres vías: extracción de características de las regiones de piel, análisis de textura y descriptores de forma de la imagen. Los tres tipos de evaluación del contenido de las imágenes son utilizados para entrenar tres clasificadores con máquinas de soporte vectorial (SVM). Los resultados de clasificación son unidos en un ensamble realizado por un metaclasificador por medio de la siguiente política: si al menos uno de los tres clasificadores concluye que la imagen es pornográfica, entonces la imagen es considerada como tal. Al evaluar todas las imágenes contenidas en una página web, se utiliza la siguiente política: si la página web presenta un porcentaje de imágenes pornográficas superior al 30%, entonces la página es considerada como pornográfica. La implementación realizada es evaluada sobre un conjunto de 5000 páginas web diversas, obteniendo una exactitud del 84.6 % en el reconocimiento de contenido pornográfico a través del contenido de las imágenes.

Palabras claves: *Clasificación de páginas web pornográficas, máquinas de soporte vectorial, detección de pornografía, aprendizaje supervisado, espacios de color.*

* Estudiantes Ingeniería de Sistemas, Grupo de Investigación en Simulación de Comportamiento de Sistemas - SICOSIS, Departamento de Ingeniería de Sistemas, Universidad de Antioquia, Medellín - Colombia {williamceballos,slest281}@gmail.com

† Magister en Ingeniería Informática, Grupo de Investigación, Desarrollo y Aplicación en Telecomunicaciones e Informática – GIDATI, Universidad Pontificia Bolivariana, Medellín - Colombia, anaisaoviedo@gmail.com

‡ Se concede autorización para copiar gratuitamente parte o todo el material publicado en la *Revista Colombiana de Computación* siempre y cuando las copias no sean usadas para fines comerciales, y que se especifique que la copia se realiza con el consentimiento de la *Revista Colombiana de Computación*.

Abstract

The World Wide Web, or web, is an information access and search logic system available on the Internet whose informative units are web pages. The web has facilitated the publication of big amount of information accessible from anywhere in the world; however, part of this content such as pornography is regarded inappropriate for some users.

To contribute to the pornography filtering on web, this paper proposes the development of a web pages classifier based on the evaluation of the images present in the webpage content. The images evaluation is done in three ways: features extraction of skin regions, texture analysis and by the shape descriptors of the image. The three types of the images content evaluation are used to train three classifiers with Support Vector Machines (SVM). The results of the SVM classification are put together in an assembly made by a metaclassifier through the following policy: if at least one of the classifiers finds that the image is pornographic, then the image is regarded as such. When assessing all the images contained in a webpage, the next policy is applied: if the webpage present a percentage above 30%, then the webpage is regarded as pornographic. The implementation done is evaluated on a set of 5000 web pages with some information kinds, getting an accuracy of 84.6% in the recognition of pornographic content through the content of the images.

Keywords: *Pornographic web pages classification, Support Vector Machines, supervised learning, colour spaces.*

1 Introducción

La web es una herramienta valiosa, que usada con responsabilidad se convierte en un instrumento educativo para niños y adultos. Sin embargo, en la web se puede acceder a información perjudicial como la pornografía, la cual afecta a menores de edad y a instituciones educativas y empresariales. Los padres de familia, educadores y empresarios necesitan un mecanismo para evitar que los niños y empleados accedan a material pornográfico a través de páginas web [10].

La información pornográfica puede ser identificada como el conjunto de materiales que muestran los órganos genitales o actos sexuales; la exhibición de dicho conjunto, normalmente tiene como objetivo la masturbación o excitación [12]. Teniendo en cuenta los prejuicios causados por el acceso a este tipo de contenidos, en este trabajo se espera aportar al reconocimiento de contenidos inapropiados en la web. Dada la gran cantidad de imágenes presentes en las páginas web con contenido pornográfico, en este trabajo se desarrolla un clasificador basado en el contenido de las imágenes.

La detección de pornografía en las imágenes es un área altamente estudiada en la actualidad y con grandes desafíos. El filtro de imágenes de Google por ejemplo, reporta una eficacia del 50% en la detección de imágenes pornográficas [23].

En este trabajo se presenta la implementación de un clasificador de páginas web pornográficas basado en el contenido de las imágenes. La detección de pornografía en las imágenes se hace por medio de la detección de regiones de piel por espacios de color, análisis de textura y descriptores de formas de los objetos presentes en las imágenes. Con las características resultantes de estos procesos de detección se entrenan tres clasificadores con Máquinas de Soporte Vectorial, conformando un metaclasificador de imágenes. El objetivo del metaclasificador es detectar contenido pornográfico en las imágenes.

El resto del artículo esta organizado de la siguiente manera: la sección 2 describe el proceso de detección de imágenes pornográficas, la sección 3 describe las Máquinas de Soporte Vectorial, la sección 4 presenta una breve revisión literaria en detección de imágenes y sistemas de clasificación de páginas web pornográficas, la sección 5 describe el método implementado, la sección 6 presenta la evaluación del prototipo construido y la sección 7 presenta las conclusiones y trabajos futuros.

2 Reconocimiento de Imágenes Pornográficas

En la detección de imágenes pornográficas, la evaluación del contenido de las imágenes es la base para el correcto reconocimiento de pornografía [29]. En esta sección se describen algunas de las técnicas en el reconocimiento de imágenes pornográficas. Una mayor fuente de consulta de técnicas en el reconocimiento de imágenes se puede encontrar en [18].

2.1 Detección de Piel

El primer paso en la evaluación de contenido pornográfico es la detección de piel [21], la cual juega un papel importante en una amplia gama de aplicaciones de procesamiento de imágenes que van desde la detección de rostros, localización de rostros y análisis de gestos [17]. Se ha demostrado que existe una fuerte correlación positiva entre el porcentaje de piel dentro de una imagen y su contenido pornográfico [22].

Para realizar la detección de piel, existen tres métodos que pueden funcionar individualmente o que se pueden combinar para obtener una mayor precisión. Estos métodos son: detección de piel por espacios de color, detección de piel por color y detección de piel por textura [18].

En general, las imágenes digitales representan la información de los píxeles por medio del espacio de color RGB (Red, Green, Blue). En este espacio de color los niveles de rojo, verde y azul claro son combinados para producir varios colores [18]. No obstante, el espacio de color RGB por sí solo no es confiable en la identificación de píxeles de color piel, ya que se debe tener en cuenta no sólo el color, sino también la luminancia [8].

La detección de piel comúnmente se desarrolla en tres pasos [14]: transformación del color del píxel a un espacio de color no RGB, disminución del componente de iluminancia del color piel, y clasificación del modelado de la distribución del color piel. Para este propósito se utiliza la combinación de los espacios de color YIQ, YUV y HSV como lo proponen en [31].

El espacio de color se puede describir como las diferentes maneras de representar o almacenar matemáticamente los colores [18]. A continuación se describen los espacios de color mencionados anteriormente.

2.1.1 Espacio de Color YUV

La piel humana tiene un rango restringido de tonos y es algo saturada [31]. YUV es un espacio de color ortogonal que permite reducir la redundancia presente en los canales de color RGB y representa el color con componentes estadísticamente independientes, tan independientes como sea posible. Diferentes investigaciones han determinado que este espacio de color es una elección favorable en la detección de piel, ya que los componentes de luminancia y crominancia están separados explícitamente [17]. De acuerdo a lo propuesto por [31], los valores de R, G y B son transformados al espacio de color YUV por medio de la ecuación (1).

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \bullet \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \text{ donde } 0 \leq R \leq 1, 0 \leq G \leq 1, 0 \leq B \leq 1 \quad (1)$$

Los componentes U y V codifican la cromaticidad de la imagen. La saturación (Ch) y el tono (θ) se pueden obtener por medio de la transformación a este espacio de color por medio de las ecuaciones (2) y (3) presentadas en [8].

$$Ch = \sqrt{|U|^2 + |V|^2} \quad (2)$$

$$\theta = \tan^{-1} (|V|/|U|) \text{ g } (180/3.14) \quad (3)$$

La saturación Ch se define como el modo de U y V y la tonalidad θ se define como el ángulo del vector en el espacio de color YUV .

2.1.2 Espacio de Color YIQ

La importancia de éste modelo en el procesamiento de imágenes es que la luminancia (Y) y la información del color (I y Q) están desacopladas, ya que la luminancia es proporcional a la cantidad de luz percibida por el ojo. Así, la importancia de este desacoplamiento radica en que el componente de luminancia de una imagen puede procesarse sin afectar su contenido cromático [9]. Los valores RGB son transformados a los valores del espacio de color YIQ de acuerdo a lo planteado por [3] por medio de la ecuación (4).

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & -0.312 \end{pmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \text{ donde } 0 \leq R \leq 1, 0 \leq G \leq 1, 0 \leq B \leq 1 \quad (4)$$

I es el eje rojo-naranja, Q es aproximadamente ortogonal a I . El menor valor de I significa menor valor azul-verde y mayor valor de amarillo [6]. El espacio YIQ puede en parte, librarse de la correlación de los componentes rojo, verde y azul en una imagen. La transformación lineal necesita menor tiempo de cómputo que las no lineales, lo cual hace que el espacio YIQ sea más preferible para sistemas no lineales. [3].

2.1.3 Espacio de Color HSV

El espacio de color HSV separa la intensidad (luminancia) de la información del color (cromaticidad) [27]. La transformación de RGB a HSV es no lineal, pero reversible. El tono (H) representa el componente espectral dominante, es decir el color en su forma pura, como en verde, rojo o amarillo. La saturación (S) se ve afectada al agregar blanco al color puro, ya que entre menos blanco, más saturado es el color. El valor (V) corresponde al brillo del color [22]. El sistema de coordenadas es cilíndrico y a menudo es representado por un subespacio definido por una pirámide invertida de 6 lados [22]. El tope de la pirámide corresponde al valor $V = 1$, con el "blanco" en el centro. El tono es medido por el ángulo alrededor del eje vertical, con rojo correspondiendo a 0° . La saturación S va desde 0 en el centro a 1 en la superficie de la pirámide. También se usa un cono invertido para denotar el subespacio en lugar de la pirámide. La transformación del espacio de color RGB a HSV según lo planteado por [31] se hace por medio de las ecuaciones (5), (6) y (7).

$$V = (R + G + B) / 3$$

$$Min = \min \{R, G, B\} \quad (5)$$

$$H = \begin{cases} 1/4 - \left(a \tan(2gR - G - B) / \sqrt{3} \right) g180 / \pi & (6) \\ \text{Cuando}(G > B) \\ 1/4 - \left(a \tan(2gR - G - B) / \sqrt{3} \right) g180 / \pi + 1/2 & (7) \\ \text{Cuando}(G \leq B) \end{cases}$$

2.2 Análisis de Textura

La textura es de gran significancia para la recuperación de imágenes basadas en el contenido [23]. El término de textura se refiere a la repetición de elementos de textura básicos llamados textones. Los textones contienen varios píxeles, cuya colocación puede ser periódica, cuasi-periódica o aleatoria. Las texturas naturales son generalmente aleatorias, mientras que las texturas artificiales son a menudo determinísticas o periódicas. La textura puede ser áspera, fina, suave, granulada, rizada, regular, irregular o lineal. Considerando la complejidad computacional y los efectos prácticos, se considerada la textura de grises en lugar de la textura de color [1].

La Matriz de Co-ocurrencia de Niveles de Grises (GLCM por sus siglas en inglés, Gray Levels Co-occurrence Matrix) utilizada por [1], es un método para medir la textura. La matriz de Co-ocurrencia es un arreglo bidimensional C en la cual, tanto las filas como las columnas representan un conjunto V de posibles tonos de gris. El valor de C (i, j) indica cuantas veces co-ocurre el valor i con j en alguna relación espacial designada.

Usualmente la relación espacial esta dada por un vector \vec{d} el cual especifica el desplazamiento entre los píxeles que tienen el valor **i** y los píxeles que tienen el valor **j**. La magnitud **r** y la dirección Θ se usan para presentar una posición (r, Θ) de un píxel. Por lo tanto, $\vec{d} = d(r, \Theta) = (dr, d\Theta)$. En consecuencia la matriz de Co-ocurrencia para una imagen **I** esta definida por:

$$C_d = \left| \{ (r, \theta) / I(r, \theta) = i \ \& \ I(dr, d\theta) = j \} \right| \quad (8)$$

Las características numéricas que pueden ser utilizadas para representar la textura más compacta son calculadas de la matriz de Co-ocurrencia. Algunas de las características estándar según [1], que se utilizan para indexar el vector de características correspondiente a textura, son definidas en las ecuaciones (9), (10), (11) y (12).

$$\text{Energía} = \sum_i \sum_j N_d^2(i, j) \quad (9)$$

$$\text{Contraste} = \sum_i \sum_j (i - j)^2 N_d(i, j) \quad (10)$$

$$\text{Correlación} = \frac{\sum_i \sum_j (i - \mu_i)(j - \mu_j) N_d(i, j)}{\sigma_i \sigma_j} \quad (11)$$

$$\text{Entropía} = - \sum_i \sum_j N_d(i, j) \log_2 N_d(i, j) \quad (12)$$

Donde μ_i , μ_j son las medidas y σ_i , σ_j las desviaciones estándar de las filas y las columnas.

Complementario a las características estándar, otras dos características tenidas en cuenta en la solución son el Segundo Momento Angular y el Momento de Diferencia Inversa, las cuales según [11], se definen en las ecuaciones (13) y (14).

$$\text{Segundo Momento Angular} = \sum_i \sum_j \{N_d(i, j)\}^2 \quad (13)$$

$$\text{Momento de Diferencia Inversa} = \sum_i \sum_j \frac{1}{1 + (i - j)^2} N_d(i, j) \quad (14)$$

2.3 Descriptores de Forma

Las imágenes pueden ser clasificadas con base a la forma de su contenido utilizando diferentes tipos de descriptores de forma [19].

El objetivo de los descriptores de forma es caracterizar la forma de un objeto. Un buen descriptor de forma debe reducir al mínimo la varianza dentro de la misma clase y maximizar la diferencia entre clases y ser insensible al ruido [30].

De las muchas medidas heurísticas de forma que se han propuesto en la literatura, la mayoría presenta un buen desempeño sólo en situaciones concretas, lo cual es una consecuencia típica del mapeo degenerado que puede resultar de la descripción de formas en términos de características (lo que implica que muchas formas pueden generar los mismos valores de características) [4].

Todas las formas con la misma área, generan el mismo valor de característica, a pesar de su gran variabilidad en forma. Debido a tal comportamiento, esta clase de características, ha sido a menudo criticada en la literatura. Sin embargo, es importante señalar que incluso algunas simples medidas de forma pueden ser aplicadas con éxito a situaciones específicas, aunque es difícil elegir cuál es el conjunto de características más adecuado para cada aplicación [4]

No existe una herramienta clara que permita determinar cuáles descriptores de forma obtendrían un mejor desempeño en cuanto a la clasificación de pornografía, por tanto, a continuación se presentan los descriptores más comunes encontrados en la literatura [25], representados en las ecuaciones (15), (16), (17), (18), (19), (20) y (21).

$$\text{Factor de forma: } 4\pi * \text{área} / \sqrt{\text{perímetro}} \quad (15)$$

$$\text{Redondez: } 4\pi * \text{área} / \pi \sqrt{e_{\text{jemayor}}} \quad (16)$$

$$\text{Compacidad: } \sqrt{(4/\pi) * \text{área} / e_{\text{jemayor}}} \quad (17)$$

$$\text{Relación de Aspecto: } e_{\text{jeprincipal}} / e_{\text{jemenor}} \quad (18)$$

$$\text{Diámetro efectivo: } \sqrt{(\text{área} / \pi) * 2} \quad (19)$$

$$\text{Solidez: } \text{área} / \text{áreaconvexa} \quad (20)$$

$$\text{Convexidad: } \text{perímetroConvexo} / \text{perímetro} \quad (21)$$

3 Métodos de Clasificación

Después de revisar el marco teórico relacionado con la detección de pornografía en las imágenes, se describe en esta sección uno de los métodos con mejores resultados en la clasificación de documentos. El marco teórico de éste método permitirá abordar la revisión de la literatura, ya que es ampliamente citado en diversas investigaciones.

Las técnicas de clasificación de páginas web han sido exploradas por medio de algoritmos de aprendizaje supervisado aplicando métodos de

clasificación de documentos. La clasificación de documentos es una tarea subjetiva que depende del criterio de un experto humano o artificial quien decide cuando clasificar un documento bajo una clase [26]. En este caso la tarea se asigna a un experto artificial como las máquinas de soporte vectorial (Support Vector Machine SVM).

Las SVM son un grupo de métodos de aprendizaje supervisado que pueden ser aplicados en problemas de clasificación o regresión [13]. Diferentes investigaciones han reportado que en tareas de clasificación, las SVM tienen un excelente desempeño [16]. Las SVM aplican técnicas lineales a problemas no lineales. Las SVM mapean un espacio de representación de características n -dimensional a un espacio de más alta dimensión, donde se puede construir un clasificador lineal mediante la definición de un hiperplano óptimo de separación de los documentos que pertenecen a la clase, de los que no pertenecen a ella. La separación de documentos por medio de un hiperplano se ejemplifica en la figura 1, donde los datos del problema original definen un espacio bidimensional y no son separables linealmente, pero al transformarlos en un espacio tridimensional ya son separables mediante un plano. El mapeo de los datos de entrada a un espacio de características de mayor dimensión se realiza por medio de una función kernel (por ejemplo: lineal, polinomial, función de base radial, sigmoideal, entre otras).

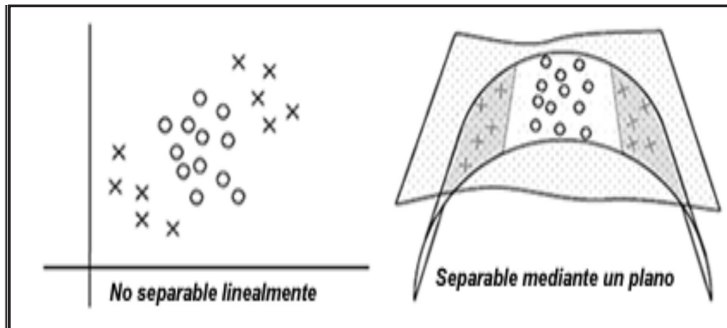


Fig. 1 Ejemplo de separación por hiperplano en SVM, tomada de [16].

Los documentos más próximos al hiperplano óptimo de separación son conocidos como muestras críticas o vectores soporte. Los vectores de soporte son los que participan en la definición del hiperplano. La búsqueda de la mínima distancia que maximice la separación de las clases se realiza a través de todas las superficies en el espacio n -dimensional, conocidas como superficies de decisión, las cuales son determinadas por vectores de soporte [16].

La evaluación de este tipo de métodos de aprendizaje supervisado, es realizada comúnmente por medio de las métricas precisión (precision), cobertura (recall), media armónica entre precisión y cobertura y exactitud. La precisión es una estimación del porcentaje de los documentos que son clasificados dentro de una clase c_i que realmente pertenecen a ella y la cobertura es una estimación del porcentaje de los documentos que pertenecen a una clase c_i que son clasificados dentro de ella [26]. Finalmente, la exactitud es la razón entre el número de documentos clasificados correctamente y el total de documentos [26].

4 Revisión Literaria

Numerosas investigaciones se han realizado para la detección de imágenes pornográficas. James Ze Wang [28], por ejemplo, utilizan ondeletas de *Daubechies* y características de histogramas de color para encontrar imágenes pornográficas. Por su parte en [24], para cada imagen se extraen características dependientes del color, características independientes del color y se aplica detección de rostros; todas estas características posteriormente son llevadas a una Máquina de Soporte Vectorial para su clasificación.

La mayoría de investigaciones en el área se han enfocado en detectar piel en el contenido de las imágenes. Forsyth y Fleck [7], por ejemplo, diseñaron un software para detectar la presencia de desnudos en una imagen analizando la estructura de las regiones de piel, dichas regiones de piel son marcadas utilizando propiedades de color y textura. Las regiones de piel identificadas son ingresadas a un agrupador especializado, el cual intenta detectar una figura humana usando limitaciones geométricas. Si el agrupador encuentra una estructura humana, el sistema informa la presencia de una imagen con desnudos [7].

Cao et al. [1] proponen un esquema de aprendizaje para detectar personas desnudas en una imagen, utilizando un vector de aprendizaje de cuantización para construir varios clasificadores basados en características de bajo nivel, tales como histogramas de color, de textura y de la forma de las regiones de piel. Finalmente utiliza un algoritmo AdaBoost para combinar los clasificadores y clasificar las imágenes.

Lee et al. [20] utiliza un espacio de color YCbCr para la detección de piel. Las características que extrae de la imagen son de tres tipos. Primero, de área, que se utilizan para cuantificar el tamaño del área examinada; segundo, de forma, que se utilizan para cuantificar la forma

en que el perfil del área examinada es similar a un cuerpo desnudo y tercero, características de ubicación, que se utiliza para cuantificar si el área examinada esta cerca al centro de la imagen.

Dentro de los espacios de color mas utilizados en sistemas de detección de piel se hallan espacios como HSV usado por [5] y [2], pero también están los espacios de color YIQ y YUV que son espacios de color ortogonales que reducen la redundancia presente en los canales de color RGB y representan el color con componentes estadísticamente independientes.

En [3] se construye un sistema para detectar imágenes pornográficas llevando a cabo tres pasos. En el paso 1, se desarrolla un modelo de color piel basado en los espacios de color YUV y YIQ para detectar las regiones de piel. En el paso 2, se utilizan bordes de Sobel y un filtro Gabor para eliminar imágenes que son erróneamente detectadas por el modelo de color piel. En el paso 3, se calcula el vector de coherencia de color y la información del histograma de color. Finalmente, se utiliza una Máquina de Soporte Vectorial para discriminar las imágenes pornográficas de las que no lo son.

En el trabajo de Zhu y Zhou [31] se presenta un modulo de detección de piel basado en los espacios de color YIQ, YUV y HSV, además de un modelo de textura con el cual busca separar los objetos que tienen color similar a piel y mejorar la exactitud en la detección de piel. Posteriormente las características calculadas son llevadas a una Máquina de Soporte Vectorial para su clasificación. En la parte experimental utilizaron 10 grupos y para medir la eficiencia y la exactitud se hace uso de criterios de desempeño general, P (Precision), R (Recall) y se define E (Error), obteniendo los siguientes resultados: $P = 88.89\%$, $R = 84.6\%$ y $E = 2.34\%$. Sin embargo, en [15] se encontró que un 94.4% de imágenes censurables podría ser detectado utilizando sólo los espacios de color YIQ y YUV.

5 Método Propuesto

La representación gráfica del método propuesto para la clasificación de páginas web se presenta en la SE figura 2.

El clasificador de páginas web recibe el código de una página y la salida es la evaluación indicando si el contenido es pornográfico o no. Internamente el clasificador está conformado por un extractor de imágenes, un metaclasificador y un evaluador.

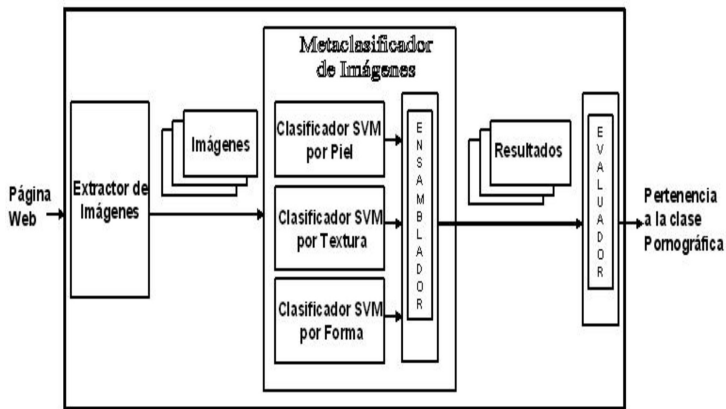


Fig. 2 Clasificador de Páginas Web

A continuación se describen los módulos internos del clasificador.

5.1 Extractor

El extractor descarga las imágenes contenidas en la página a través de las etiquetas “img” del código HTML.

5.2 Metaclasificador de Imágenes

El Metaclasificador analiza cada una de las imágenes de forma independiente por medio de tres clasificadores previamente entrenados según sus características de piel, textura y forma. Se denomina metaclasificador dado que está compuesto por los tres clasificadores mencionados anteriormente, los cuales son construidos en una etapa previa a la clasificación de las páginas web. A continuación se presenta una descripción de los módulos internos del Metaclasificador.

5.2.1 Clasificador SVM por Piel

Las características extraídas en el módulo de detección de piel, una vez las regiones de piel fueron identificadas, se organizan en un vector, el cual representa la imagen inicial que se tenía como dato de entrada en el proceso de detección de pornografía, es decir, este vector es la forma de indexar las imágenes para el clasificador.

Como un paso previo al proceso de detección de pornografía, el clasificador SVM es entrenado con diferentes tipos de imágenes

preclasificadas en las clases pornográfica y no pornográfica. Este proceso de entrenamiento permite construir el hiperplano de separación de imágenes de este tipo de contenido.

Al entrar en el módulo de clasificación con SVM por piel, se determina si el vector ingresado pertenece a la clase pornográfica, dependiendo del lado del hiperplano del cual este ubicado. Finalmente, la salida del proceso es un valor arrojado por el clasificador que indica la pertenencia a la clase pornográfica.

5.2.2 Clasificador SVM por Textura

Por su parte, en el proceso de extracción de características para textura, se toma la imagen y se transforma a niveles de grises, a partir de la cual se calcula la matriz de co-ocurrencia para obtener las características correspondientes a la textura. Con estas características se forma el vector con el cual se indexa la imagen para llevarla al clasificador de textura.

El clasificador SVM por textura también es entrenado previamente con diferentes tipos de imágenes etiquetadas en las clases pornográfica y no pornográfica. Este proceso de entrenamiento permite construir el hiperplano de separación de imágenes de este tipo de contenido. Al entrar en el módulo de clasificación con SVM, se determina si el vector ingresado pertenece a la clase pornográfica, dependiendo del lado del hiperplano del cual este ubicado. Finalmente, la salida del proceso es un valor arrojado por el clasificador que indica la pertenencia a la clase pornográfica.

5.2.3 Clasificador SVM por Formas

El objetivo de los descriptores de forma es caracterizar la “forma” de un objeto. Un buen descriptor de forma debe reducir al mínimo la varianza dentro de la misma clase y maximizar la diferencia entre clases y ser insensible al ruido [30]. Tomando como base esta afirmación optamos por utilizar descriptores de forma en las imágenes para nuestro tercer clasificador.

Como ocurrió en los anteriores clasificadores, previo a hacer pruebas de clasificación es necesario entrenar el clasificador SVM por formas con imágenes de diferentes tipos etiquetadas en las clases pornográfica y no pornográfica. Este proceso de entrenamiento permite construir el hiperplano de separación de imágenes de este tipo de contenido.

Al entrar en el módulo de clasificación con SVM, se determina si el vector ingresado pertenece a la clase pornográfica, dependiendo del

lado del hiperplano del cual este ubicado. Finalmente, la salida del proceso es un valor arrojado por el clasificador que indica la pertenencia a la clase pornográfica.

5.3 Ensamblador

El ensamblador recibe los resultados de los tres clasificadores y es el encargado de determinar si una imagen corresponde o no a la clase pornográfica. Para este fin se determina una política en la que una imagen es considerada como pornográfica siempre y cuando al menos uno de los clasificadores arroja como resultado que la imagen es pornográfica.

Este módulo constituye la salida final del metaclasificador, cuyo resultado es entregado al evaluador que dará el veredicto final para la clasificación de páginas web.

5.4 Evaluador

El evaluador se encarga de decidir cuando una página web es pornográfica. La entrada de este módulo está conformada por los resultados de la evaluación de cada una de las imágenes contenidas en la página.

Para determinar si una página tiene contenido pornográfica se implementa la siguiente política: si al menos 30% de las imágenes son pornográfica, entonces la página se considera pornográfica también. Para realizar dicha evaluación, este módulo espera el resultado de todas las imágenes presentes en la página.

6 Evaluación del Prototipo

Para evaluar el método propuesto para la clasificación de páginas web pornográficas se construyó un prototipo, el cual es evaluado en un diseño experimental con las siguientes características.

Los datos experimentales son tomados del *Open Directory Project - ODP* <<http://www.dmoz.com>>. En el ODP existen mas de 4,6 millones de recursos, como páginas web, grupos de noticias, servidores FTP, entre otros. Los recursos organizados en más de 590,000 categorías, son mantenidos por casi 81,462 editores humanos de manera gratuita.

Para evaluar el prototipo construido se seleccionó una muestra aleatoria de 5000 páginas web, siguiendo la distribución de los datos existentes en Internet publicada en el sitio web <<http://www.internet-filter-review.com>>. La distribución de las páginas es presentada en la figura 3.

Distribución de paginas Web

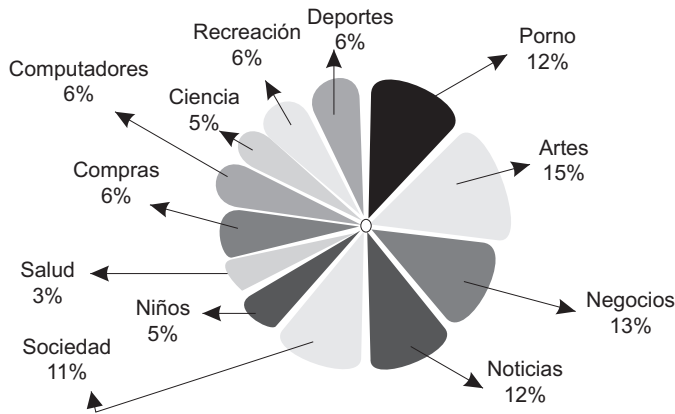


Fig. 3. Distribución de tipos de información.

Los resultados de la evaluación del clasificador son presentados en la tabla 1.

Clasificación			
<i>Precision</i> <i>p</i>	<i>Recall</i> <i>r</i>	<i>Exactitud</i> <i>e</i>	<i>Media Armónica</i> <i>f1</i>
0,430	0,858	0,850	0,573

Tabla 1. Resultados evaluación del clasificador de páginas web.

De los datos de la tabla se puede concluir que el 85.8% (cobertura r) de las páginas pornográficas fueron reconocidas, dejando de identificar el 14.20% de las páginas pornográficas. También se puede concluir que 43.0% (precisión p) de las páginas reconocidas como pornográficas realmente lo eran, lo cual indica que se tiene un 57.0% de falsas alarmas.

El valor de la exactitud nos indica que se clasificaron correctamente el 86% del total de páginas de prueba.

Un resultado importante que se debe comentar es el valor de la precisión, ya que se está clasificando un número importante de páginas como pornográficas cuando no lo son. La clasificación de las páginas depende de la correcta clasificación de las imágenes que esta contiene. Los resultados de la precisión son atribuibles a algunas características de las imágenes como brillo, luminosidad, calidad de la imagen, entre otros que obstaculizan una clara lectura de su contenido.

Adicionalmente, las páginas no pornográficas pueden ser mal clasificadas si contienen imágenes de personas en primer plano en donde el color predominante sea el color de la piel, un ejemplo claro son las imágenes de rostro humano las cuales son mal clasificadas por el clasificador de piel.

Otro factor influyente sobre todos los resultados en general radica en los datos de prueba, ya que se observó que cuando la página está construida con frames no se puede acceder a las imágenes del frame principal donde se encuentran las imágenes, lo que produce la falla del clasificador de páginas.

7 Conclusiones

En este trabajo se presentó un método para la clasificación de páginas web pornográficas basado en el contenido de las imágenes. Para evaluar el clasificador propuesto se construyó un prototipo, el cual fue probado en un conjunto de 5.000 páginas web. En el experimento se obtuvieron los siguientes resultados: el 85.8% de las páginas pornográficas fueron reconocidas, dejando de identificar el 14.20% de las páginas pornográficas; el 43.0% de las páginas reconocidas como pornográficas realmente lo eran, lo cual indica que se tiene un 57.0% de falsas alarmas.

El alto porcentaje de las falsas alarmas es justificado por problemas en el clasificador de imágenes basado en la piel, ya que puede reconocer como pornográficas imágenes que no lo son.

Como trabajos futuros se propone implementar un detector de rostros humanos en las imágenes, ya que cuando estas se presentan en primer plano en una imagen, dado su alto contenido de píxeles de color piel tienden a ser clasificadas erróneamente como pornográficas.

Además se propone afinar los rangos de las reglas de las transformaciones para eliminar el brillo y lograr detectar con mayor exactitud la piel. Igualmente, para los clasificadores de textura y forma

se propone experimentar con nuevas características con el fin de establecer mejores espacios de separación entre las clases.

Referencias

- [1] L. L. Cao, X. L. Li, N. H. Yu, Z. K. Liu, "Naked People Retrieval Based on Adaboost Learning," International Conference on Machine learning and Cybernetics vol. 2, pp. 1133 - 1138, 2002.
- [2] Y. Chan, R. Harvey, D. Smith, "Building systems to block pornography," In Challenge of Image Retrieval, BCS Electronic Workshops in Computing series, pp. 34-40, 1999.
- [3] H. D. Cheng, X. H. Jiang, Y. Sun, J. Wang, "Color image segmentation: advance and prospects," Pattern Recognition 34, pp. 2259–2281, 2001.
- [4] L. F. Costa and R. Cesar, "Shape Analysis and Classification: Theory and Practice," CRC Press, cap 6, 2001.
- [5] C. Cusano, C. Brambilla, R. Schettini, G. Ciocca, "On the Detection of pornographic digital images," Visual Communications and Image Processing (VCIP2003), vol. 5150, pp. 2105-2113, 2003.
- [6] L. Duan, G. Cui, W. Gao, H. Zhang, "Adult image detection method base-on skin colour model and support vector machine," Asian Conference on Computer Vision, pp. 797-800, 2002
- [7] M. Fleck, D. A. Forsyth, C. Bregler, "Finding Naked People," Proceedings European Conference on Computer Vision and Pattern Recognition, pp. 593-602, 1996
- [8] M. R. Girgis, T. M. Mahmoud, Tarek, "An Approach to Image Extraction and Accurate Skin Detection from Web Pages," International Journal of Computer Science and Engineering, Vol 1, no. 2, pp. 88-96, 2007
- [9] R. C. González & R.E. Woods, "Digital Image Processing," Prentice Hall, 2002.
- [10] M. Hammami, Y. Chahir and L. Chen, "Webguard: Web based adult content detection and filtering system," in Proceedings of the International Conference on Web Intelligence (WI'2003), pp. 574–578, 2003.
- [11] R.M. Haralick and K. Shantnugam and I. Dinstein. "Textural features for image classification", IEEE Trans. SMC, vol 3 no 6, p610-6211. 97
- [12] L. Hunt, The invention of pornography, Zone Books, New

- York, 1993.
- [13] O. Ivanciuc “Applications of Support Vector Machines in Chemistry,” In: Reviews in Computational Chemistry, Vol 23, Eds.:K.B. Lipkowitz and T.R. Cundari. Wiley-VCH, Weinheim, pp. 291-400, 2007
 - [14] S. Jayaram, S. Schmugge, M.C. Shin, L.V. Tsap, “Effect of Colour space Transformation, the Illuminance Component, and Colour Modelling on Skin Detection,” CVPR, pp. 813-818, 2004.
 - [15] F. Jiao, W. Gao, L. Duan, G. Cui, “Detecting Adult Image using Multiple Features,” International Conferences on Info-tech & Info-net 2001 (ICII 2001), vol. 3, pp. 378-383, 2001.
 - [16] T. Joachims, “Text categorization with support vector machines: learning with many relevant features”, en Proceedings of ECML-98, 10th European Conference on ML, number 1398 in Lecture Notes in Computer Science, Springer Verlag, Heidelberg, DE, 1998, pp.137-142
 - [17] P. Kakumanu, S. Makrogiannis, N. Bourbakis, “A survey of skin-colour modelling and detection methods,” Pattern Recognition, vol. 40, pp. 1106-1122, 2007.
 - [18] W. Kelly, A. Donnellan, D. Molloy, "A review of skin detection techniques for objectionable images," the 7th Information Technology and Telecommunication Conference, pp. 40- 49, 2007.
 - [19] I. Kunttu, L. Lepistö, J. Rauhamaa, A. Visa, “Multiscale Fourier Descriptor for Shape Classification,” Proceedings of 12th International Conference on Image Analysis and Processing, pp. 536-541, 2003
 - [20] J.S. Lee, Y.M. Kuo, P.C. Chung, E.L.Chen, “Naked image detection based on adaptive and extensible skin colour model,” Pattern Recognition, vol. 40, Issue 8, pp.2261–2270, 2006.
 - [21] K.M. Liang, S.D. Scott, M. Waqas, “Detecting pornographic images,” Asian Conference on Computer Vision ACCV, pp. 497-502, 2004
 - [22] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, A. Yamada, “Color and texture descriptors,” IEEE Trans. CSVT 11, vol 6, pp. 703–715, 2001.
 - [23] J. S. Payne, L. Hepplewhite, T. J. Stonham, “Perceptually based metrics for the evaluation of textural image retrieval methods,” Multimedia Computing and Systems, IEEE International Conference on, vol 2, pp. 793-797, 1999.
 - [24] H. Rowley, Y. Jing, S. Baluja, “Large scale image-based adult-

- content filtering”, VISAPP2006, pp. 290-296.
- [25] J. C. Russ, “The Image Processing Handbook,” 3rd ed., CRC Press, Boca Raton, FL
 - [26] F. Sebastiani, Machine Learning in Automated Text Categorization. En: ACM Computing Surveys, p.p 1-47, 2002.
 - [27] S. Sural, G. Qian, S. Pramanik, “Segmentation and histogram generation using the HSV color space for image retrieval,” Proceedings of the IEEE International Conference on Image Processing, IEEE, Rochester, pp. 589–592, 2002.
 - [28] J. Z. Wang, J. Li, G. Wiederhold, O. Firschein, “System for Screening Objectionable Images,” Computer Communications Journal, vol. 21, no. 15, pp. 1355-1360, Elsevier, 1998
 - [29] S. Xuanjing, W. Wei, Q. Qingji, “The Filtering of Internet Images Based on Detecting Erotogenic-part”, Natural Computation, 2007. ICNC 2007. Third International Conference on, Volume 2, 24-27 Aug. 2007 Page(s):732–736.
 - [30] Q. F. Zheng, M. J. Zhang, W. Q. Wang, “Shape-based Adult Images Detection,” Third International Conference on Image and Graphics ICIG2004, Hong Kong, China, pp. 150-153, 2004.
 - [31] H. Zhu, S. Zhou, J. Wang, Z. Yin, “An algorithm of pornographic image detection,” in Fourth International Conference on Image and Graphics, Sichuan, China, pp. 81-84, 2007.