

# **Segmentación y Clasificación de Números Provenientes de Tarjetas de Registro de Asistencia**

Isidro Moctezuma Cantorán, Remedios Fabián Velasco, José Francisco Delgado Orta

Universidad del Mar, Campus Puerto Escondido, Ciudad Universitaria s/n, Puerto Escondido, San Pedro Mixtepec, Oaxaca, México.

**Resumen:** En este trabajo se presentan los resultados de una aplicación del método de perfiles para la segmentación y reconocimiento de números en imágenes digitales, y el método de correlación para la clasificación de números, los cuales son obtenidos de las tarjetas de asistencia de una institución de educación pública.

**Palabras claves:** Tarjetas de registro, reconocimiento de caracteres, procesamiento de imágenes.

## **I. Introducción:**

Las técnicas de procesamiento de imágenes han facilitado las tareas para el reconocimiento de objetos, utilizando sistemas de cómputo en problemas específicos, como de identificación de rostros, placas de automóviles (Cárdenas-Hidalgo, y otros 2009), huellas dactilares, traducción de documentos antiguos (Toscano-Medina, y otros 2005), detección de tumores en mamografías, entre otras aplicaciones (Faúndez-Zanuy 2001).

Entre de las técnicas de procesamiento de imágenes se encuentran las de Reconocimiento óptico de caracteres (OCR), cuyas aplicaciones están dirigidas al tratamiento automático de textos (Arlandis Navarro 2007); dentro del OCR destaca la técnica de los perfiles que se enfocan o permiten hacer la segmentación de las formas básicas de las imágenes (Faúndez-Zanuy 2001), las cuales corresponden a números, letras, símbolos u otras formas, principal objeto de estudio de los problemas antes mencionados.

Dentro de las tareas administrativas en una institución de educación pública está el control de asistencias. Tarea que involucra la utilización de dispositivos de entrada como lectores de huella digital, lectores de tarjetas, cámaras de reconocimiento de rostros o iris, entre otros, donde el procesamiento de imágenes adquiere suma importancia en la autenticación de los trabajadores o empleados para el registro de sus entradas y salidas.

En ocasiones el manejo de tarjetas en adjunción con grandes cantidades de registros conlleva a un control lento, tardado y complejo. Para realizar una automatización o mejoramiento se requiere de un proceso de conversión de los datos provenientes de formatos físicos de tarjetas de asistencia a imágenes digitales; donde se pueden aplicar las técnicas de procesamiento de imágenes (Arlandis Navarro 2007), para la detección de los números que representan los horarios de las entradas y salidas, con el fin de determinar las incidencias de acuerdo a las normas de la organización.

## **II. Problemática**

La universidad del Mar es una institución de educación superior pública, ubicada al sureste de la República Mexicana, cuenta con tres campus: Huatulco, Puerto Ángel y Puerto Escondido, en los cuales laboran alrededor de 400 trabajadores entre administrativos, académicos y de mantenimiento. De acuerdo al reglamento de personal [(Universidad del Mar 2004)] los trabajadores deben cumplir con una jornada laboral preestablecida según su contrato, el incumplimiento de esta norma deriva una incidencia. Para llevar el registro y control de asistencias, que incluye entrada y salida, se utilizan 3 mecanismos: primero, el trabajador tiene asignada una tarjeta de asistencia, en la cual debe registrar el tiempo utilizando en un reloj checador; segundo, el

De esta forma, de los tres mecanismos el de mayor utilización es el de tarjetas de asistencia con el 90% de los trabajadores. Habiendo una persona asignada para detectar incidencias, tarea que realiza en forma manual usando el formato de la figura 1. El formato se compone de celdas con la información de identificación y el tipo de trabajador en la parte superior. En la sección inferior del formato, las celdas contienen los registros de las entradas y salidas de cada trabajador obtenidas a partir del reloj checador. De esta forma, la persona asignada revisa los datos de cada una de las celdas, para comparar posteriormente, si dichos datos corresponden a una incidencia de acuerdo a su horario establecido.

[illegible]

No.		NOMBRE		2																																																																																																	
DEPARTAMENTO		PERIODO DE PAGO																																																																																																			
Cance la do																																																																																																					
<table border="1"> <thead> <tr> <th colspan="2">TIEMPO REGULAR</th> <th colspan="2"></th> <th colspan="2">TIEMPO EXTRA</th> </tr> <tr> <th>IN</th> <th>OUT</th> <th>IN</th> <th>OUT</th> <th></th> <th></th> </tr> </thead> <tbody> <tr><td>01:30</td><td>01:30</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:30</td><td>01:30</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:30</td><td>01:30</td><td></td><td></td><td></td><td></td></tr> <tr><td></td><td></td><td></td><td></td><td></td><td></td></tr> <tr><td>01:00</td><td>01:00</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:00</td><td>01:00</td><td>01:50</td><td>01:50</td><td>01:00</td><td>01:00</td></tr> <tr><td>01:00</td><td>01:00</td><td>01:00</td><td>01:00</td><td>01:00</td><td>01:00</td></tr> <tr><td></td><td></td><td></td><td></td><td></td><td></td></tr> <tr><td>01:00</td><td>01:00</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:00</td><td>01:00</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:04</td><td>01:04</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:04</td><td>01:04</td><td></td><td></td><td></td><td></td></tr> <tr><td>01:04</td><td>01:04</td><td></td><td></td><td>01:04</td><td>01:04</td></tr> <tr><td>01:05</td><td>01:05</td><td>01:05</td><td>01:05</td><td>01:05</td><td>01:05</td></tr> </tbody> </table>						TIEMPO REGULAR				TIEMPO EXTRA		IN	OUT	IN	OUT			01:30	01:30					01:30	01:30					01:30	01:30											01:00	01:00					01:00	01:00	01:50	01:50	01:00	01:00	01:00	01:00	01:00	01:00	01:00	01:00							01:00	01:00					01:00	01:00					01:04	01:04					01:04	01:04					01:04	01:04			01:04	01:04	01:05	01:05	01:05	01:05	01:05	01:05
TIEMPO REGULAR				TIEMPO EXTRA																																																																																																	
IN	OUT	IN	OUT																																																																																																		
01:30	01:30																																																																																																				
01:30	01:30																																																																																																				
01:30	01:30																																																																																																				
01:00	01:00																																																																																																				
01:00	01:00	01:50	01:50	01:00	01:00																																																																																																
01:00	01:00	01:00	01:00	01:00	01:00																																																																																																
01:00	01:00																																																																																																				
01:00	01:00																																																																																																				
01:04	01:04																																																																																																				
01:04	01:04																																																																																																				
01:04	01:04			01:04	01:04																																																																																																
01:05	01:05	01:05	01:05	01:05	01:05																																																																																																
<p>DECLARAR QUE LA PRESENTE TABLA DE PAGO MUESTRA PERSONALMENTE POR M. A LAS HORAS DE ENTRADA Y SALIDA Y POR LO TANTO CORRESPONDE AL RESCATE DE M. RESCATEADO.</p>																																																																																																					
FIRMA DEL EMPLEADO																																																																																																					

b) Reverso

Dada la periodicidad de las tareas descritas y el tiempo que consume el personal asignado para la ejecución de las mismas en el procedimiento manual, el presente trabajo propone una aplicación con un método de detección de números en imágenes, correspondientes a las tarjetas de registro de asistencia, extrayendo la información de las celdas, verificando las incidencias que corresponden a la jornada laboral y generando automáticamente un reporte de incidencias relacionadas a la jornada laboral de los trabajadores.

Los trabajos de reconocimiento de caracteres, se enfocan en procesos de segmentación y detección de símbolos alfanuméricos. (Cárdenas-Hidalgo, y otros 2009) Implementan en un sistema de reconocimiento automático de placas de automóviles, capaz de procesar imágenes,

obteniendo información característica de las mismas y almacenándola en una base de datos, desde una interfaz de usuario elaborada en Matlab. Para ello utilizan fotografías de las placas de 150 a 210 píxeles y de 50 a 110, las cuales deben tomarse horizontalmente.

Para ello, se transforma la imagen original eliminando información no relevante y definiendo una función para calcular el umbral óptimo de cada fotografía particular. A esta imagen, se le aplica una transformación morfológica "bottom-hat" (Leymarie y Levine, 1988), a través de una operación "closing" (dilatación y erosión), para posteriormente restar la imagen obtenida de la imagen original, resaltando las placas y los bordes. Después se realiza la operación "opening" (erosión y luego dilatación) con un elemento estructurante vertical, basándose en la altura de los caracteres de la placa para, dilatarla vertical y horizontalmente expandirla. Una vez expandida, se binariza y recorta la zona de la placa con un umbral de 120 para elementos que producen ruido y se binariza para depurar más la imagen.

Por último, se aplica una propiedad de región para que solo las áreas mayores a 1.5% de total queden presentes. Recortando cada carácter por separado utilizando la función bwlable, la cual divide la imagen en regiones y determina el área de cada región (abarcando el área ocupada por el carácter), para recortarlo con el comando incrop, obteniendo cada carácter en una variable por separado. A cada carácter obtenido, se le redimensiona a una imagen de (24\*42) para estandarizar los caracteres y poder realizar el reconocimiento de la letra y del número. A partir de esto, empieza el desarrollo del OCR, para ello se define una base de datos de las 24 letras del alfabeto en mayúsculas y los signos del 0 al 9, empleando imágenes de 21\*24 píxeles, binarizados, que se cargan en matlab mediante variables, y luego se crea una matriz llamada alfabeto que contiene las letras del alfabeto y una matriz llamada numeral que contiene todos los dígitos. Por último, utilizan la correlación como método para reconocimiento de carácter hallado en la placa con uno presente en la base de datos, dando como resultado un coeficiente que presenta la igualdad entre dos matrices.

Otros enfoques utilizan algoritmos de clasificación basados en redes neuronales artificiales para el reconocimiento de texto. Por ejemplo, (Velásquez, Sossa-Azuela y Levachkine 2002), implementan redes neuronales artificiales (RNA) de tipo multicapa para el reconocimiento de caracteres alfanuméricos obtenidos desde mapas impresos, los cuales representan la forma más común de información geográfica. Para ello identifican provee información, como contornos, límites, redes hidráulicas, carreteras, etc., acompañado de leyendas de diferentes tipos, letras, número, símbolos y de diferentes tamaños e inclinaciones.

Dada la cantidad de recursos que requiere el procesamiento de la imagen, se producen mapas vectoriales; éstos se obtienen a partir de mapas ráster, separando cada mapa en sus capas, donde cada capa representa los caracteres alfanuméricos (letras y números). La identificación de estos, tiene un alto grado de dificultad debido a que las letras y números pueden ser de diferente fuente, tamaño e inclinación, por lo cual, se construyó un algoritmo bastante robusto para aislar conjuntos de píxeles y una vez aislados, utilizar clasificadores especialmente diseñados para reconocer cada carácter.

Este reconocimiento se hace con una RNA Backpropagation (Rumelhart, Hinton y Williams, 1988), un paradigma para ajustar los pesos de las redes, consiste en ajustar primero los pesos que conectan a las dos últimas capas, considerando las diferencias de los valores esperados y los valores obtenidos, a continuación de van propagando hacia atrás a las demás pares de capas ocultas, mediante un delta, hasta llegar a las que conectan la capa de entrada con la primera capa oculta. La red de este trabajo se entrenó con diferentes tipos de fuentes, tamaños e inclinaciones.

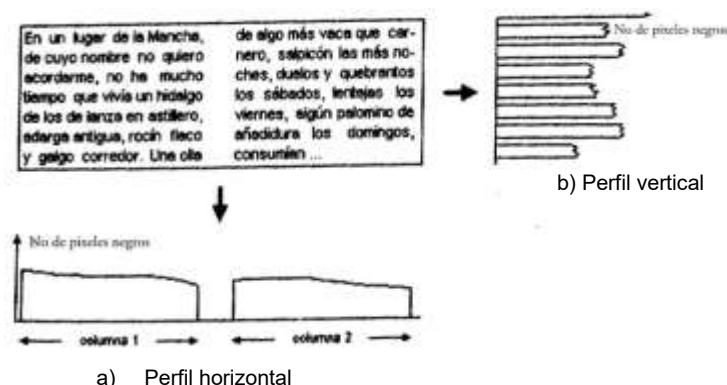
El proceso de aprendizaje de la red, hace una selección de los caracteres para tener una muestra representativa, usando 9 fuentes de 18 tamaños diferentes (desde los 8 puntos hasta los 25 puntos) y las inclinaciones abarcan 90° con incrementos de 10°. El aprendizaje consiste en convertir todos los caracteres en archivos bmp (Microsoft Windows Bitmap), para introducirlos a la

red neuronal como plantillas binarias. La conversión de los archivos bmp se realizó con MATLAB, separando los caracteres y formando mallas cuadradas para alimentar a la RNA, así como caracteres especiales y falsos positivos que fueron agregados como “ruido” para la pruebas de la red. Para el diseño de las RNA se consideraron tres tipos de caracteres: minúsculas, mayúsculas y números, y cada red contempla dos tamaños y tres ángulos de inclinación; con esos aspectos formaron 81 RNA provenientes de los 9 grupos de tamaños, los tres grupos de inclinaciones y los tres grupos de caracteres.

Las pruebas de la RNA, consistieron en seleccionar varios tipos de caracteres, utilizando las tres redes que contemplan todas las inclinaciones, si no se identifica el carácter se hace otro intento con su esqueleto, con el objetivo de identificar todos los caracteres independientemente de que sean falsos positivos. Esta implementación reconoció el 95.51% de los caracteres alfanuméricos provenientes de mapas raster escaneadas a color, un porcentaje alto a pesar de que las redes no fueron entrenadas con muestras reales.

Una de las aplicaciones con mayor utilización desde su aparición es el reconocimiento de caracteres ópticos (OCR, Optical Character Recognition) (). El objetivo es traducir una imagen escaneada de un texto manuscrito o impreso a un archivo de caracteres. Las aplicaciones comunes de OCR han sido la identificación automática de matrículas de vehículos, lectura de códigos postales de cartas, talones, facturas, etc. Lógicamente el problema es más sencillo en el caso de texto impreso que en texto manuscrito, puesto que en el primer caso el número de fuentes posibles es pequeño (habitualmente letra normal, cursiva y negrita), y las realizaciones de una misma letra siempre son iguales. En cambio, en texto manuscrito la variabilidad es mucho mayor, es difícil segmentar entre letras (separarlas). Aunque, la existencia de separaciones entre los elementos de la imagen, facilita el procesamiento para la segmentación y se utiliza para aprovechar la información disponible para su reconocimiento.

Tal es el caso del método de perfiles, el cual, utiliza la información que proporcionan los píxeles de la imagen para segmentar las letras de un texto impreso. Para ello, se enfoca en el cálculo de un perfil sobre un eje determinado, operación que consiste en proyectar todos los píxeles sobre el eje en cuestión. De esta forma, se pueden obtener perfiles horizontales y verticales de una imagen en 2 dimensiones. La obtención del perfil horizontal consiste en sumar para cada columna de la imagen el número de píxeles negros. La figura 2.a, muestra el perfil horizontal (obtenido mediante proyección vertical) de dos párrafos. En ella, se distinguen los márgenes izquierdo y derecho del texto, y la separación entre columnas.



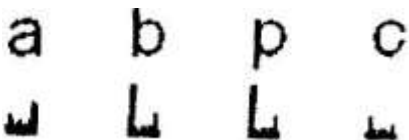
**Figura 2.** Localización de columnas y líneas en un fragmento de texto.

La proyección vertical de la figura 2.b., permite identificar las líneas de texto y el espaciado interlineal. Una vez obtenidas las líneas, la proyección horizontal dentro de cada línea proporciona los límites de los diferentes caracteres. Por tanto, el resultado final es la posición de los caracteres

que componen el texto. El método contempla en el paso siguiente, identificar cada uno de los caracteres leídos de la imagen con los símbolos posibles del alfabeto. Para ello existen diversas alternativas, de las que se distinguen dos: a) La aplicación del método de los perfiles para distinguir caracteres, y b) el reconocimiento óptico de caracteres mediante una red neuronal.

De esta forma, el uso de perfiles se utiliza para detectar párrafos, líneas y caracteres, así como distinguir letras. Por ejemplo, la figura 3 muestra los perfiles horizontales de las letras a, p, b y c. En ella se puede observar que existen letras con perfiles horizontales iguales (p y b) y, por tanto, un único perfil no es suficiente para distinguir todos los caracteres. En la práctica, se realizan perfiles sobre dos o más ejes, y la comparación se realiza con los perfiles correspondientes almacenados de cada uno de los caracteres del abecedario.

La valoración de los resultados obtenidos mediante un determinado OCR puede especificarse mediante una matriz de confusión, en la que idealmente todos los elementos de la diagonal principal deberían valer el 100% en el caso de no haber errores. Si se producen errores, la matriz de confusión nos indica el porcentaje de error de las distintas combinaciones (número de veces que se ha confundido cada carácter con el resto).



**Figura 3.** Perfiles horizontales de las letras a, b, p y c.

Una vez capturado y binarizado el texto es necesario delimitar los caracteres de forma individual. En el caso de texto manuscrito, un gran número de caracteres están conectados y, por tanto, resulta difícil separarlos. De hecho, un OCR sobre caracteres aislados puede ser mejor que una persona, pero no se ha conseguido superar a las personas en el reconocimiento de caracteres conectados. Dependiendo del tipo de datos presentados a la red neuronal que actuará como clasificador de la entrada a uno de los caracteres de salida posibles, hay dos grandes grupos de reconocedores:

a) Mapa de bits de los caracteres. Se escala cada uno de los caracteres a reconocer a un tamaño fijo (por ejemplo, 16x16 píxeles) y se presenta a la red el mapa de bits normalizado. Si durante el entrenamiento de la red se le presentan un número suficiente de realizaciones distintas de cada uno de los caracteres a reconocer, la red poseerá una buena capacidad de generalización y podrá reconocer entradas no utilizadas en la fase de entrenamiento.

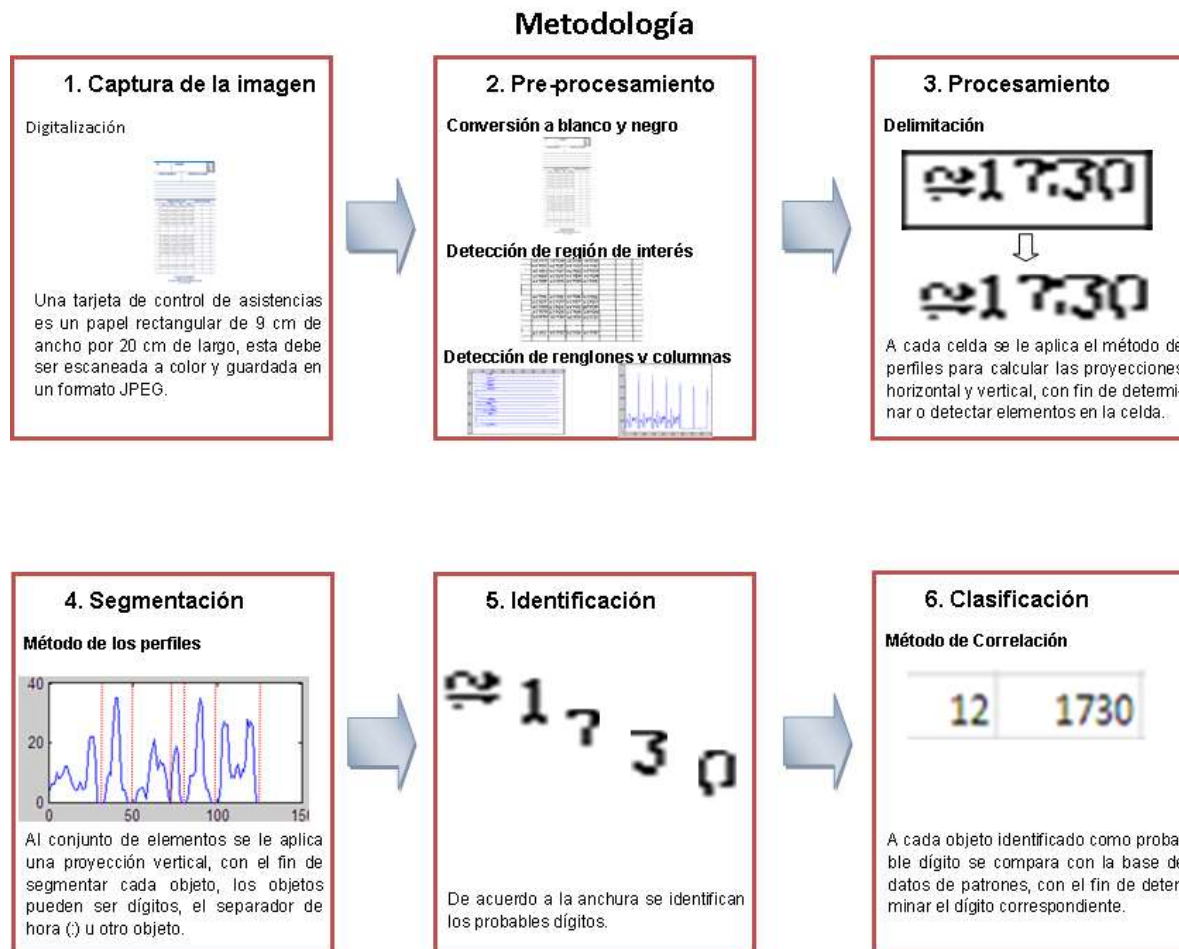
b) Característica de los caracteres. Para reducir el tamaño de la red, y por tanto, la complejidad de su entrenamiento, se extrae la información relevante de los caracteres mediante un extractor de características, tales como el número de agujeros, características direccionales, puntos terminales, etc.

Sin embargo, el extractor de características definido como una red neuronal suele ser una opción no viable, porque existe una gran variabilidad entre los mapas de bits de diversas realizaciones de un mismo carácter, mientras que las características presentan una variación menor que posibilita el uso de una red de menor complejidad mediante el segundo esquema, el cual muestra una opción viable para la tarea de reconocimiento.

#### **IV. Metodología de solución.**

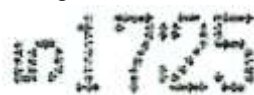
El trabajo propuesto, formula una metodología basada en las técnicas para el pre-procesamiento de la imagen, el método de perfiles para la segmentación de caracteres y el método de correlación para la clasificación de los caracteres. Estos se implementaron en Matlab, utilizando el paquete de

funciones diseñadas para trabajar con técnicas para el procesamiento de imágenes. De esta forma, la secuencia de tareas para la solución del problema se describe en la figura 4.



**Figura 4.** Metodología empleada.

La metodología empleada consta de 6 bloques principales, iniciando desde la tarjeta física y generando como salida un archivo en excel de acuerdo con el objetivo del OCR descrito en (Faúndez-Zanuy 2001). El primer bloque consiste en capturar una imagen de la tarjeta física de la figura 1, a partir de un escáner con resolución de 300 a 600 puntos por pulgada (dot per inch o d.p.i.) y almacenarla en un archivo JPG. Después en el segundo bloque, se preprocesa la imagen usando binarización, calculando el área de interés (las celdas correspondientes a los horarios de entrada y salida), y aplicando el método de perfiles para la detección de los renglones y columnas sobre las líneas del formato, obteniendo con el cálculo de los perfiles, los delimitadores de las líneas del formato. Posteriormente, en el bloque tres, se eliminan de la imagen los delimitadores de renglones y columnas de la imagen original, filtrando solo el contenido de la celda en una imagen intermedia, como la que se muestra en la figura 5.



**Figura 5.** Imagen de una celda de la tarjeta de asistencia.

A ésta imagen, se le hace el cálculo de los perfiles con su proyección horizontal y vertical de los números de la celda para obtener la información de la hora y el día en que se registró la entrada o salida en la tarjeta. La información de la imagen de la figura 5 presenta dos tipos de números, aquellos con orientación horizontal corresponden a la hora de entrada o salida del trabajador; mientras que los números que aparecen en orientación vertical, corresponden al día del mes en que el trabajador hizo el chequeo de su entrada o salida. Una vez calculados los perfiles, se procede a segmentar la imagen en la tarea 4, aplicando una proyección horizontal para buscar el umbral de separación de los dígitos. Después, en el bloque 5 se identifican los dígitos uno a uno verticalmente de acuerdo al formato de horas, y en grupos de uno o dos dígitos para el formato del día. Se procesan en imágenes intermedias para filtrarlas como entrada al clasificador del bloque 6, el cual usa el método de correlación para hacer el reconocimiento.

Las tareas de conversión y segmentación de las imágenes emplean los comandos de MatLab (Cárdenas-Hidalgo, y otros 2009). De esta forma, se formó la base de patrones en formato de imágenes bmp, de acuerdo a (Velásquez, Sossa-Azueta y Levachkine 2002) con 31 para la clasificación de las fechas (grupos de uno o dos dígitos) y 10 para los dígitos de las horas (números naturales del 0 al 9).

## V. Pruebas y resultados

La clasificación obtenida con la metodología propuesta para una tarjeta de asistencia de un periodo de 15 o 16 días mostrada en la figura 6, se almacena en un archivo de salida en formato Excel mostrado en la figura 7.

De acuerdo a las pruebas realizadas en una computadora portátil Dell XPS con procesador Intel Core 2 duo a 2 GHz, con 4GB de memoria RAM, con la metodología propuesta se clasificó el 96.7% de los dígitos de manera correcta con un error del 1.89%.

Figura 6. Imagen la tarjeta de asistencia.

Fecha	IN	OUT	IN	OUT	Tiempo Extra
1	2127	2127	2127	2127	0
2	2127	2127	2127	2128	0
3	2128	2128	2129	2128	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	2129	2129	0	0	0
8	2130	2130	0	0	0
9	2130	2130	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0
12	0	0	0	0	0
13	1331	1331	0	0	0
14	1331	1331	0	0	0
15	1332	1332	0	0	0

Figura 7. Resultado de la detección en Excel.

## **VI. Conclusiones**

En conclusión, con esta metodología se logró hacer una clasificación correcta del 96.7% de los dígitos correspondientes a las tarjetas de asistencia, con un error de 1.89%. Para mejor funcionamiento se recomienda usar imágenes de tarjetas que no estén manchadas, rayadas o mutiladas, debido a que éstas características no fueron consideradas para las pruebas.

Con el resultado de esta metodología se corroboró que el método de los perfiles y el método de correlación, son excelentes herramientas para segmentar y clasificar respectivamente, ya que arrojaron resultados cercanos al 100% esperado.

## **Bibliografía**

Arlandis Navarro, Joaquim. «OCR. Sistemas de reconocimiento óptico de caracteres.» Actualidad TIC (Revista del Instituto Tecnológico de Informática), nº 13 (Noviembre 2007): 9-11.

Cárdenas-Hidalgo, Paul, José Alfredo Flores-Vargas, Jaime López-Zavaleta, y Pablo Martínez-Moreno. Diseño de sistema de reconocimiento de placas utilizando Matlab. Tesina, Escuela Superior de Ingeniería Mecánica y Eléctrica unidad Culhuacan, Instituto Politécnico Nacional, México D. F.: Instituto Politécnico Nacional, 2009.

Chong-Wah, Ngo, y Chan Chi-Kwong. «Video text detection and segmentation for optical character recognition.» Editado por University of Hong Kong Department of Computer Science. Multimedia Systems (Springer-Verlag), nº 10 (Diciembre 2005): 261-271.

Faúndez-Zanuy, Marcos. Tratamiento digital de voz e imagen y aplicación a la multimedia. México: Alfaomega, 2001.

Leymarie, F. y M. D. Levine. Curvature morphology. Technical Report. TR-CIM-88-26, Computer Vision and Robotics Laboratory, McGill University, Montreal, Quebec Canada, 1988.

Rumelhart, D., G.E. Hinton y J. R. Williams. Learning Internal Representations by Error-Propagation. En D.E. Rumelhart y J.L. McClelland, eds., Parallel Distributed Processing, Volumen 1, capítulo 8. (1988). pp. 675-695.

Toscano-Medina, Karina, Mariko Nakano-Miyatake, Gabriel Sánchez-Pérez, Héctor M. Pérez-Meana, y Makoto Yasuhara. «Reconocimiento de caracteres manuscritos usando la función spline natural.» Editado por Sección de Estudios de Posgrado e Investigación. Científica 9, nº 3 (2005): 143-154.

Universidad del Mar. Reglamento interior de trabajo. Puerto Ángel, Oaxaca, 17 de Agosto de 2004.

Velásquez, Aurelio, Juan Humberto Sossa-Azuela, y Serguei Levachkine. «Reconocimiento eficiente de caracteres alfanuméricos provenientes de mapas ráster por medio de clasificadores neuronales.» Computación y Sistemas (Instituto Politécnico Nacional) 6, nº 001 (2002): 38-50.