

# Instituto Tecnológico de Costa Rica

## Escuela de Ingeniería en Computación

### Bachillerato en Ingeniería en Computación

**Curso:** IC-8046 Introducción al Reconocimiento de Patrones

**Profesor:** Ing. Eduardo A. Canessa Montero, M.Sc.

**Semestre:** II Semestre 2020

**Tarea de Investigación No. 4:** Reconocedor estadístico básico de dígitos escritos a mano con el uso de histogramas como vector de entrada.

**Fecha de entrega:** Lunes 16 de noviembre de 2020.

#### I. INSTRUCCIONES GENERALES.

- Esta tarea debe de ser desarrollada de forma grupal en los mismos grupos de trabajo del proyecto final.
- La solución de esta tarea deberá de ser subida al TEC Digital en la sección de evaluación correspondiente a la Tarea 04 antes de las 11:59pm de la fecha de entrega indicada en este enunciado.
- Entregables: Se deberá de entregar un archivo comprimido con toda la información generada (imágenes, código fuente y artículo)
- Se requiere de la utilización de una herramienta para procesamiento digital de imágenes.
- En esta tarea se programará un reconocedor estadístico básico utilizando como datos de entrada del modelo el vector con la información del histograma de una imagen binarizada.

#### II. INDICACIONES.

1. La primera parte de este trabajo será desarrollada como Laboratorio 05, en el cual se tendrá que preparar lo siguiente:

- Cada miembro del equipo de trabajo escribirá 30 repeticiones (especímenes) de cada dígito del sistema numérico decimal, esta información deberá de ser entregada (escaneada o fotografiada) como parte del laboratorio.
- Generar un programa que extraiga de forma individual cada dígito escrito a mano en imágenes separadas, de forma tal que se tenga una muestra general de imágenes de cada vocal. Cada una de las imágenes individuales deberá de pasar por un preprocesamiento que limpie las imágenes y las deje todas en la misma posición, lo más alineadas posibles y solo con la información necesaria (sin ruido), y todas con las mismas dimensiones.
- Generar un programa que extraiga el histograma (con la información tanto horizontal como vertical) con el conteo de los píxeles que forma parte de la letra, el ancho del histograma para considerar esta información deberá de ser de 4 píxeles. Se deberá de presentar visualmente mediante un gráfico de barras.
- Como evidencia del desarrollo del Laboratorio 05 deberá de subir al rubro correspondiente un archivo ZIP con lo obtenido en este punto (imágenes y código fuente).

2. Utilizando el 70 % de los especímenes obtenidos en el punto anterior y seleccionados de forma aleatoria, desarrolle un programa que extraiga el modelo estadístico (valor promedio y varianza) del histograma de cada una de las 5 vocales para utilizarlo como reconocedor.

3. Utilizando el 30 % restante de los especímenes obtenidos en el punto 1. y el modelo estadístico del punto anterior, desarrolle un programa que presentando la imagen de la vocal muestre de qué vocal se trata
4. Tabule la tasa de aciertos obtenida al evaluar todos los especímenes utilizados en el punto anterior para cada uno de los dígitos en decimal.
5. Investigue sobre reconocedores de caracteres escritos existentes en la comunidad de software libre, y distintos modelos conceptuales de implementación de sistemas para el reconocimiento de números escritos a mano, de los mismos, de acuerdo a su investigación seleccione dos y describa como funcionan. Guarde esta información y todo lo realizado en esta tarea, ya que será utilizado posteriormente en el desarrollo del Primer Proyecto del curso
6. De acuerdo con la investigación del punto 5, formule 3 hipótesis relacionadas con los resultados comparativos de efectividad del método desarrollado en esta tarea comparados con cada uno de los dos métodos seleccionados, para un total de 6 hipótesis.
7. Debe de generar un reporte en formato de artículo científico de acuerdo al formato *IEEE Transactions*. En este reporte se debe de incluir una descripción detallada de los dos reconocedores implementados, todos los resultados obtenidos y tabulados, el análisis de resultados del punto anterior y al final del artículo un resumen con las 3 conclusiones más importantes que se deriven de presente trabajo.
9. Este artículo debe de realizarse en idioma inglés, lo cual tendrá un valor de 10 % del presente trabajo.