

IMPLEMENTACIÓN DE RECONOCIMIENTO DE CARACTERES EN LA CAPTURA DE DATOS PARA LA CALIBRACIÓN DE EQUIPOS SIN INTERFAZ DE COMUNICACIÓN

Cristy K. Sánchez S¹, Isaac E. Ruíz A¹.

Centro Nacional de Metrología de Panamá, CENAMEP AIP.
Rep. de Panamá • Ancón • Ciudad del Saber • Edificio 206 • Código Postal: 0843-01353
+ (507) 517-3156 csanchez@cenamep.org.pa

RESUMEN: La transformación digital se ha convertido una necesidad que aporta valor agregado a los laboratorios de calibración; dicho esto, el Centro Nacional de Metrología de Panamá, tomó la decisión de automatizar el proceso de adquisición de datos para la calibración de equipos con pantalla LCD y sin interfaz de comunicación, los servicios de estos dispositivos son solicitados regularmente por parte de la industria. Actualmente la calibración de estos equipos se realiza manualmente lo que implica la escritura de los datos en folios, el cual se digitaliza como evidencia y se realiza la transcripción en hojas de cálculo para el análisis y obtención de los resultados que se presentan en los certificados de calibración. Todo este proceso aumenta el tiempo de espera del certificado de calibración.

1. INTRODUCCIÓN

Los programas de interpretación de imágenes por lo general son de uso bajo licenciamiento comercial, lo cual derivar en un costo de implementación, por ende el Centro Nacional de Metrología de Panamá, busca obtener un programa adecuado a los procedimientos del laboratorio, de licenciamiento libre y que permita la escalabilidad y la rápida adaptación a otros proyectos, y con ello poder conocer la incertidumbre asociada a las operaciones. Por tal motivo se desarrolló el programa en lenguaje Python [1], con la capacidad de utilizar diferentes técnicas o librerías para el reconocimiento de caracteres, implementando Tesseract-OCR, este es un motor de reconocimiento de caracteres de software libre. OCR-Reconocimiento de Caracteres Ópticos es una aplicación de reconocimiento de patrones que corresponden a textos contenidos en imágenes [3]; con esta implementación podemos transformar las imágenes con texto a cadenas de caracteres y así procesarlo en nuestro programa.

2. DESCRIPCIÓN

Se desarrolló un programa que puede capturar e interpretar los valores numéricos que se muestran en las pantallas LCD de los equipos que no tienen una interfaz de comunicación que nos permita acceder a estos y obtener información. Para poder manipular el programa se creó una interfaz de usuario basada en librerías de herramientas gráficas QT5 [4], para organizar la metodología de la adquisición de los datos según el procedimiento interno de calibración del laboratorio [11] ("Fig. 1"). La interfaz permite la visualización del entorno de trabajo y el control de todas las operaciones de modificación de luminancia, gamma o morfológicas que se realizan sobre la imagen. Para el tratamiento del video y la modificación de los fotogramas no basamos en la librería OpenCV [5]. Los valores obtenidos son almacenados en base de datos, con la utilización de MariaDB (o MySQL) cuando existe un servidor dedicado o SQLite para el almacenamiento local.

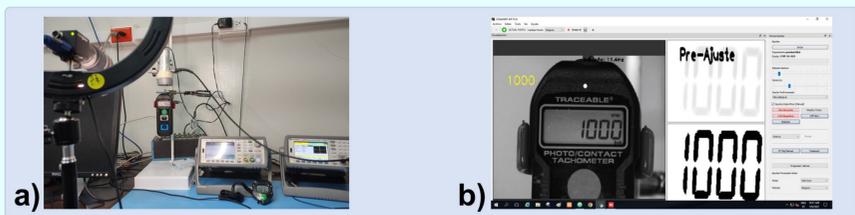


Fig. 1: a) Equipo bajo calibración (izquierda). b) Interfaz gráfica (derecha).

Durante la captura de la imagen y su procesamiento se utiliza NTP (Network Time Protocol o Protocolo de Tiempo por Red) [6], para la generación de la estampa de tiempo y obtención de la rastreabilidad a la Hora Nacional del CENAMEP AIP.

2.1 Desarrollo

El programa se creó con la idea de reproducir la toma de datos que realiza el metrólogo, para esto se desarrolló un Software que cumpliera con la lectura de los dígitos mostrados en la pantalla del equipo bajo calibración, constando de tres módulos ("Fig. 2") que interactúan entre sí de manera concurrente.

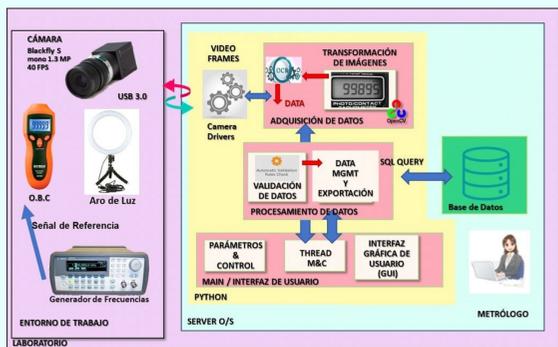


Fig. 2: Arquitectura Modular del Programa.

El primer módulo es la interfaz del usuario, aquí el responsable de la calibración (el metrólogo) realizará la creación del proyecto de toma de datos en el programa. Luego parametrizará el proyecto ("Fig. 3") según el tipo de equipo que tiene, entre las opciones incluidas están los puntos de medición o de calibración. La ventana de adquisición se refiere al tiempo en que se registrará la captura del valor y la cantidad de datos a adquirir. Para optimizar la interpretación el programa tiene otros parámetros a configurar en la barra de herramientas ("Fig.4"), selección del área de interés, el tipo de pantalla (si es transreflectiva negativa o positiva), la inserción de las cifras significativas y el número de decimales para las pantallas que el punto decimal está muy abajo y las opciones de Tesseract, donde se seleccionará el archivo de datos de entrenamiento para obtener el mejor resultado de extracción de la información.

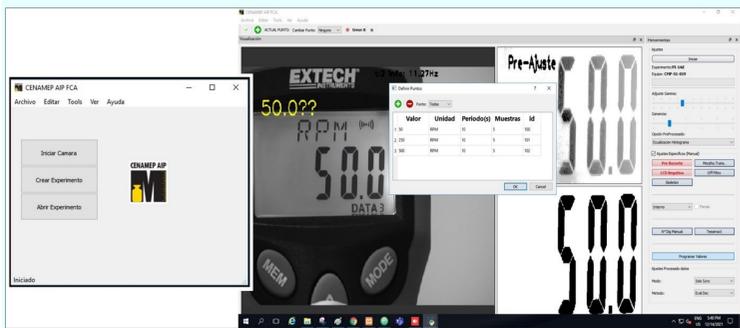


Fig. 3: Ventana inicial del programa.



Fig. 4: Panel de Herramientas configuración del programa.

El segundo módulo del programa es el encargado de realizar la validación de la data y verificar que lo obtenido cumpla con un conjunto de preceptos ("Fig. 5"), primeramente, que sea un valor numérico, que la cantidad de unidades y valores decimales sea la esperada y que el error de la medida con respecto al valor de referencia se mantenga en cierto rango. En caso de encontrar una irregularidad en el dato adquirido, el valor es rechazado y se solicita al módulo de adquisición que repita el proceso y tome un nuevo valor. Para el caso que el valor sea aceptado, este pasará a ser depositado en la base de datos.

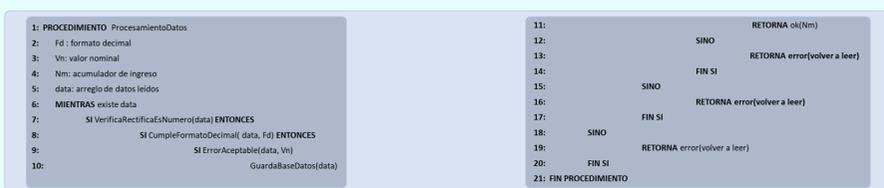


Fig. 5: Algoritmo de validación de la data.

3. RESULTADOS

Para que el programa pueda ser implementado en las calibraciones se analizaron los resultados de diferentes pruebas para conocer su comportamiento y determinar si es metrologicamente apto, siguiendo las recomendaciones del punto 7.11 de la Norma ISO/IEC 17025:2017 [7]. Entre las pruebas están el Ensayo de Aptitud y la de coincidencia y consistencia. Para el Ensayo de Aptitud se realizó una comparación entre la toma de datos automática y una toma realizada manualmente por un metrólogo, en el cual obtuvimos un resultado satisfactorio según el análisis del Error Normalizado ("Tabla 1", "Tabla 2", "Tabla 3" y "Gráfica 1").

$$En = \frac{|m_{AdqAut} - m_{AdqM}|}{\sqrt{U^2_{AdqAut} + U^2_{AdqM}}}$$

Ecuación 1: Fórmula de evaluación del Error normalizado.

Punto	Ref (RPM)	Obj (RPM)	Error (RPM)	U (RPM)
1	50.00	50.0	0.0	0.1
2	250.00	250.1	0.1	0.2
3	500.00	500.1	0.1	0.5
4	1 000.0	999.9	-0.1	0.2
5	2 500.0	2 499	-1	1
6	5 000.0	4 999	-1	1
7	10 000.0	9 999	-1	1
8	25 000.0	24 999	-1	1
9	50 000.0	49 999	-1	2
10	93 750.0	93 747	-3	2

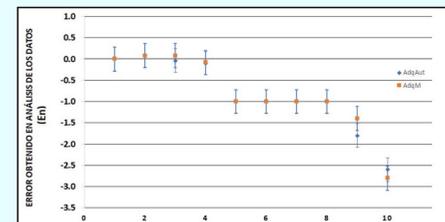
Tabla 1: Resultados obtenidos durante la toma de datos manual.

Punto	Ref (RPM)	Obj (RPM)	Error (RPM)	U (RPM)
1	50.00	50.0	0.0	0.1
2	250.00	250.1	0.0	0.4
3	500.00	500.0	0.0	0.4
4	1 000.0	999.9	-0.1	0.1
5	2 500.0	2 499	-1	1
6	5 000.0	4 999	-1	1
7	10 000.0	9 999	-1	1
8	25 000.0	24 999	-1	1
9	50 000.0	49 998	-2	2
10	93 750.0	93 747	-3	2

Tabla 2: Resultados obtenidos durante la toma de datos automática con el programa.

Punto	Ref (RPM)	Resultado (En)
1	50.00	0.0
2	250.00	0.0
3	500.00	0.2
4	1 000.0	0.1
5	2 500.0	0
6	5 000.0	0
7	10 000.0	0
8	25 000.0	0
9	50 000.0	0
10	93 750.0	0

Tabla 3: Resultados obtenidos al aplicar la fórmula del En.



Gráfica 1: Resultado del error en el método manual y el método automático.

Se realizó la prueba de análisis de coincidencia entre la imagen capturada y el valor que el programa le determinó ("Fig. 6" y "Fig. 7"), para así validar el método de conversión de las imágenes a datos cumple con el requerimiento tanto de compatibilidad con el proceso de calibración como también de aceptación según los procedimientos internos del laboratorio.

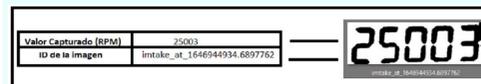


Fig. 6: Validación del programa, imagen coincidente.

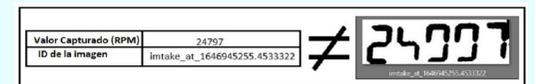


Fig. 7: Validación del programa, imagen no coincidente.

De la prueba de coincidencia y consistencia de los datos donde se analizaron 10 000 capturas obtuvimos que un total de 9 745 registros que fueron interpretados de manera correcta, mientras que 56 registros fueron incorrectos, lo que arroja un total del 99.45% de acierto global ("Tabla 4").

Frecuencia de Referencia / Frecuencia Nominal (RPM)	Ventana de Adquisición en segundos (s)			
	1	5	10	20
50.00	100%	100%	100%	100%
100.00	100%	100%	100%	100%
500.00	83%	100%	100%	100%
750.00	100%	100%	100%	100%
1 000.0	100%	100%	100%	100%
2 500.0	100%	100%	99%	100%
5 000.0	100%	100%	100%	100%
7 500.0	100%	100%	100%	100%
10 000.0	100%	98%	100%	100%
15 000.0	100%	99%	99%	100%
50 000.0	100%	100%	100%	100%
93 750.0	100%	99%	100%	100%

Tabla 4: Porcentaje de coincidencia.

Otra de las pruebas realizada fue la estabilidad del programa ante variaciones del entorno, se tomaron 10 000 (diez mil) datos de un valor fijo en la pantalla, para determinar que el proceso de adquisición no se vea afectado por elementos externos como el parpadeo de los tubos de luz fluorescentes, a una frecuencia de 60 Hz que es la de línea eléctrica de Panamá, durante esta prueba los datos se capturaron sin mayor problema, dando un 100% de acierto.

3.1 Discusión

Los resultados obtenidos en las pruebas revelan que el programa tiene similitud con la captura manual en fidelidad, aun así, es recomendable realizar optimización de la evaluación de los datos para los casos en que los dígitos numéricos en la pantalla del equipo quedan incompletos debido a la rapidez del refrescado del LCD en algunos equipos. Pese a esto el peor porcentaje de acierto fue de 83% para el punto 500.00 RPM ("Tabla 4"), con la versión original del módulo de evaluación. Sin embargo, se ha trabajado en la actualización de este módulo para lograr obtener mejores resultados. Las verificaciones aquí plasmadas para la validación del sistema se obtienen empleando dos modelos de equipos bajo calibración, aunque se verificó la funcionalidad del reconocimiento de caracteres en pantallas hasta con 7 equipos de diferentes marcas.

4. CONCLUSIONES

La presentación de este trabajo cumple con nuestra propuesta de ser una herramienta de ayuda en el Centro, que optimiza el proceso de captura de datos para equipos sin interfaz de comunicación, brindándole al metrólogo flexibilidad en la ejecución de su trabajo y mitigando la generación de posibles errores de transcripción. Los datos recabados de las pruebas, con tasas globales de errores de 0.4 % y donde el ensayo de aptitud entre el humano y el programa cumple satisfactoriamente permite seguir con la implementación y el cumplimiento del objetivo final del proyecto que es el uso del programa para los servicios regulares de calibración de tacómetros.

5. REFERENCIAS

- [1] Ozgur, C., Colliau, T., Rogers, G., & Hughes, Z. (2017). MatLab vs. Python vs. R. Journal of Data Science, 15(3), 355-371.
- [2] Smith, R. (2007, September). An overview of the Tesseract OCR engine. In Ninth international conference on document analysis and recognition (ICDAR 2007) (Vol. 2, pp. 629-633). IEEE.
- [3] Chaudhuri, A., Mandaviya, K., Badelia, P., & Ghosh, S. K. (2017). Optical character recognition systems. In *Optical Character Recognition Systems for Different Languages with Soft Computing* (pp. 9-41). Springer, Cham.
- [4] Harwani, B. M. (2018). Qt5 Python GUI Programming Cookbook: Building responsive and powerful cross-platform applications with PyQt. Packt Publishing Ltd.
- [5] Howse, J., & Minichino, J. (2020). Learning OpenCV 4 Computer Vision with Python3: Get to grips with tools, techniques, and algorithms for computer vision and machine learning. Packt Publishing Ltd.
- [6] Mills, D. L. (1989). On the accuracy and stability of clocks synchronized by the network time protocol in the internet system. ACM SIGCOMM Computer Communication Review, 20(1), 65-75.
- [7] ISO/IEC 17025:2017: <https://www.iso.org/obp/ui/#iso:std:iso-iec:17025:ed-3:v2:es>
- [8] ITU., Recomendación UIT-R TF.686-3, Glosario y definiciones de términos de tiempo y frecuencia. Unión Internacional de Telecomunicaciones. 2013.
- [9] Batagelj, V., Bojkovski, J., & Drnovšek, J. (2008). Software integration in national measurement-standards laboratories. IET science, measurement & technology, 2(2), 100-106.
- [10] Mollah, A. F., Majumder, N., Basu, S., & Nasipuri, M. (2011). Design of an optical character recognition system for camera-based handheld devices. arXiv preprint arXiv:1109.3317.
- [11] R. Solís., PE-S1-001: Procedimiento interno de Calibración de Medidores de Frecuencia de Acople Óptico, del Centro Nacional de Metrología de Panamá. Versión 6 (2021).

6. AGRADECIMIENTOS

Agradecemos al Centro Nacional de Metrología de Panamá por el apoyo que nos ha brindado al permitirnos desarrollar este proyecto y a la vez proporcionarnos todo lo necesario para llevarlo a término y al compañero Raúl F. Solís por su orientación y recomendaciones.