**Proyecto: Red Neuronal Convolucional aplicada en MNIST reconocimiento de dígitos manuscritos para la solución de CAPTCHA**

JAIRO ANGEL

SEPTIEMBRE 2024

BOOTCAMP DE INTELIGENCIA ARTIFICIAL

TALENTO TECH

**Problema a resolver**

Para permitir el acceso a diferentes servicios solicitados por ciudadanos se requiere la consulta y certificación de estados de aseguramiento, este proceso requiere la consulta a unas páginas web que con el número de identificación del ciudadano expiden el certificado de aseguramiento respectivo, este proceso se surte previa validación de un CAPTCHA específico que presenta una imagen de 5 dígitos, a los que se le agrega ruido aleatorio. Para facilitar este proceso de consulta masivo y continuo se implementan procesos de RPA (Robotic Process Automation) que permiten a diferentes entidades procesar de forma automatizada el manejo de estas certificaciones. En el proceso de automatización se requiere resolver este CAPTCHA que consiste en la determinación de un número que se presenta en una imagen con las siguientes características:



**Pertinencia**

Hay un costo humano y de tiempo de proceso al no tener una herramienta que permita la consulta y obtención de certificados de forma automatizada dada la demanda de servicios requeridos sin tener que resolver el captcha solicitada por la página.

**Estado del Arte**

No hay un api o servicio para resolver de forma automatizada el captcha especificado

Las tecnologías disponibles en la solución de este captcha compuesto por este tipo de imágenes son:

1. Redes neuronales convolucionales (CNN): las CNN han sido el enfoque más exitoso para el reconocimiento de dígitos MNIST. Técnicas como DropConnect1, redes neuronales profundas multicolumna (MCDNN)1 y redes de cápsulas (CapsNet)1 han logrado altos índices de precisión.
2. Aprendizaje en conjunto: combinación de múltiples modelos para mejorar el rendimiento. Por ejemplo, EnsNet1 y RMDL (Random Multimodel Deep Learning)1 han mostrado mejoras significativas.
3. Aumento de datos: se han utilizado técnicas como APAC (Clasificación de patrones aumentada)1 para mejorar los datos de entrenamiento, lo que conduce a una mejor generalización.

Referencias

LeCun, Y., Cortes, C., & Burges, C. J. C. (1998). [The MNIST database of handwritten digits](http://yann.lecun.com/exdb/mnist/" \t "_blank)[Link2](http://yann.lecun.com/exdb/mnist/).

Baldominos, A., Saez, Y., & Isasi, P. (2019). A Survey of Handwritten Character Recognition with MNIST and EMNIST. Applied Sciences[, 9(15), 3169](https://www.mdpi.com/2076-3417/9/15/3169" \t "_blank)[Link3](https://www.mdpi.com/2076-3417/9/15/3169).

Wikipedia. [MNIST databaseLink4](https://en.wikipedia.org/wiki/MNIST_database).

Papers With Code. [MNIST Benchmark (Image Classification)](https://paperswithcode.com/sota/image-classification-on-mnist) [MNIST Benchmark (Image Classification) | Papers With Code](https://paperswithcode.com/sota/image-classification-on-mnist) [1](https://paperswithcode.com/sota/image-classification-on-mnist).

**Ventajas Comparativas**

Permitir el acceso a los certificados con la solución de este captcha permite la automatización del proceso liberando recurso humano que debe consultar esta información. Esta automatización permite que diferentes entidades con un alto volumen de consulta puedan reducir sus costos y generar una mejor oportunidad en el trámite de los diferentes servicios que requieren este certificado.

**Marco Legal y Ético que aplica y condiciona el proyecto**

El servicio de consulta de información y certificados está disponible para este fin. Las entidades que consultan esta información deben tener procedimientos definidos y claros relacionados con La Ley 1266 de 2008, también conocida como Ley de Habeas Data, se aplica a todos los datos personales financieros, crediticios, comerciales y de servicios registrados en un banco de datos.

**Posibles riesgos y supuestos / Situaciones externas que delimitan el proyecto**

Los modelos planteados no resuelvan adecuadamente el captcha planteado incurriendo en costos de procesamiento y no obtención de los certificados requeridos

Cambio en la estrategia de la página web sobre la que se realizan las consultas.

Aceptación de la herramienta por parte de las entidades que requieren realizar las consultas.

**Objetivo**

Se propone la construcción de un servicio de “Solución de CAPTCHA Aseguramiento” que usando una Red Neuronal Convolucional entrenada con el conjunto de datos MNIST desarrollado en Python usando la estrucutra general de Red Neuronal Convolucional presentada por el profesor Anfres Felipe Escallon en el material Talento Tech Bootcamp Inteligencia Artificial (2024). Este servicio se desplegara con flask.palletsprojects para que reciba una imagen y determine el número que se presenta en esta imagen.

**Metodología y Cronograma propuesto**

Este proyecto se llevara a cabo bajo las metodologías de control de proyectos PMI y utilizando metodologías agiles y Extreme programming en la conformación de los equipos humanos para el desarrollo del software.

Se llevaran a cabo las siguientes etapas:

1. Mecanismos APIs de recepción y carga de imagen
2. Desarrollo de procesos de filtrado de Imagen
3. Proceso de determinación de dígitos y fraccionamiento de imagen
4. Construcción de Red Neuronal Convolucional entrenada con MNIST (reconocimiento de dígitos manuscritos) y validación del funcionamiento.
5. Realización de la predicción y retorno de resultados

|  |  |
| --- | --- |
| Semana | Actividad |
| 1 | Mecanismos APIs de recepción y carga de imagen |
| 2-3 | Desarrollo de procesos de filtrado de Imagen |
| 4-5 | Proceso de determinación de dígitos y fraccionamiento de imagen |
| 6-10 | Construcción de Red Neuronal Convolucional entrenada con MNIST (reconocimiento de dígitos manuscritos) |
| 11-12 | Construcción módulo de predicción y retorno de resultados, evaluación |
| 13-14 | Creación de apis para consumo por sistemas externos |