

Caracterización de fibras ópticas de un detector de partículas usando IA

Jorge Eduardo Ossa Sánchez
Estudiante de doctorado Universidad de Medellín
Doctorado en Modelación y Computación Científica

9th ComHEP: Colombian Meeting on High Energy Physics
San Juan de Pasto

5 de diciembre de 2024

Motivación

En la implementación del sistema de calibración del detector de fotones del detector lejano del proyecto DUNE y del sistema de potencia para los dispositivos CE, los cables de fibra óptica juegan un papel fundamental.

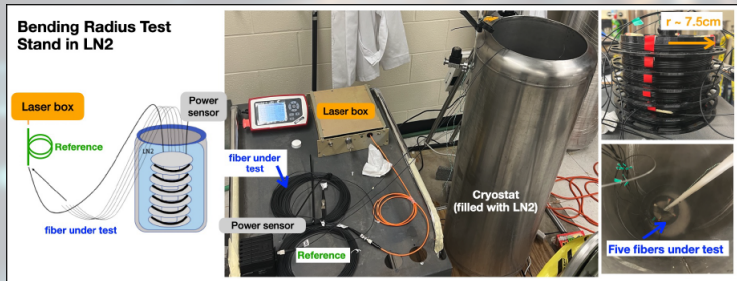
Se utilizarán segmentos de fibra óptica de hasta 25 mt sumando un aproximado de 54 km solamente para la aplicación PoF.

¿Cómo ayudar en la automatización del proceso de inspección visual de ~ 54 kilómetros de fibra óptica?

Equipo de trabajo

Phd Alex Tapia
Phd David Vanegas
Est David López
Est Jorge Ossa

Power over fiber



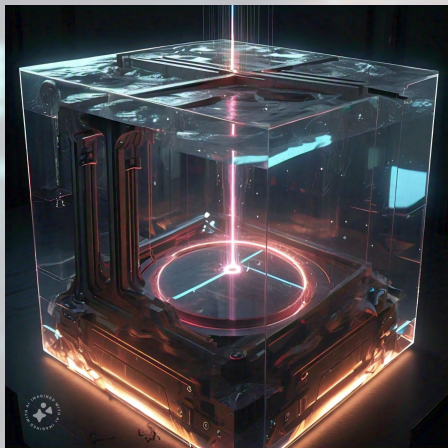
Basado en:

Characterization and Novel Application of Power Over Fiber for Electronics in a Harsh Environment

<https://arxiv.org/abs/2405.16816v1>

<https://doi.org/10.48550/arXiv.2405.16816>

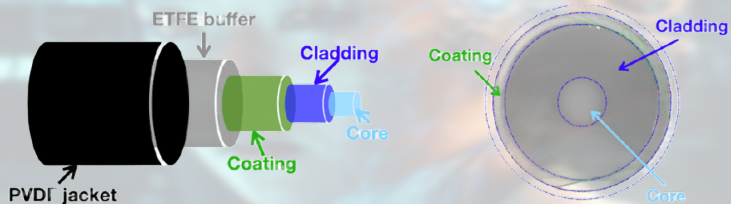
Enfoque



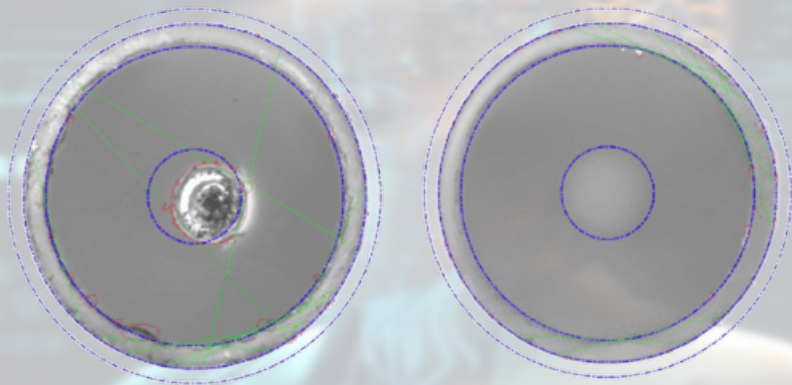
En el marco de trabajo del detector de fase única del detector lejano del proyecto DUNE:

¿Cómo ayudar a determinar si un cable de fibra óptica es apto para transmitir potencia?

Estructura de una fibra óptica

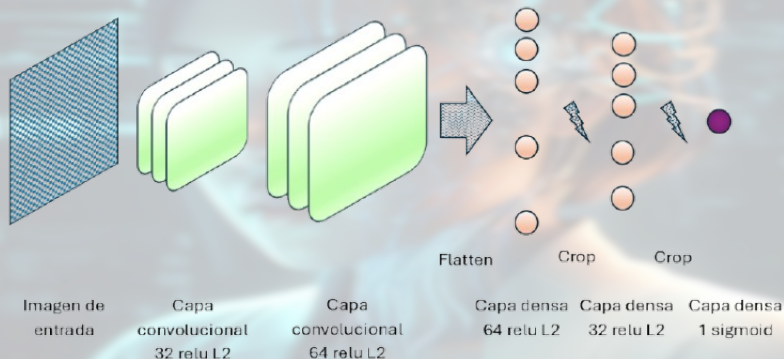


Corte transversal de una fibra óptica



¿Cómo seleccionar una fibra óptica basado en la imagen de su corte transversal?

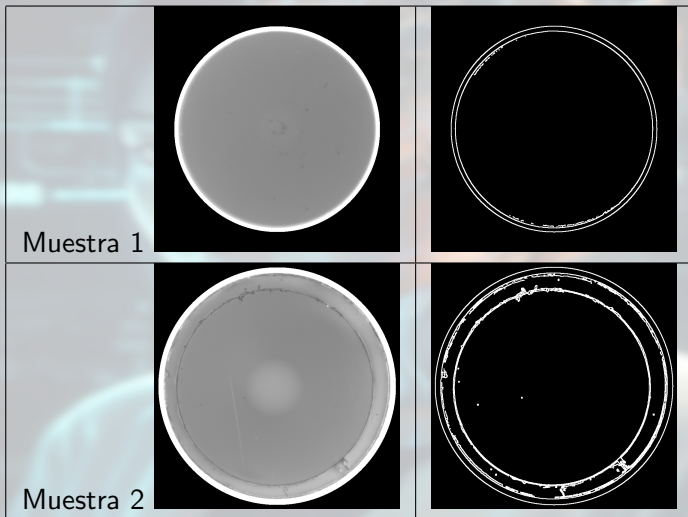
Red neuronal



El entorno de desarrollo → Python + Colab + GPU

- ▶ tensorflow
 - ▶ red neuronal
 - ▶ ImageDataGenerator
 - ▶ TensorBoard
- ▶ numpy
 - ▶ arreglos
 - ▶ matrices
- ▶ cv2
 - ▶ procesamiento de imágenes
- ▶ matplotlib
 - ▶ graficación de imágenes
- ▶ os
 - ▶ manejo de archivos
- ▶ sklearn
 - ▶ manipulación de muestras

Dificultades I



Dificultades II

- ▶ Pequeña cantidad de muestras de entrenamiento
- ▶ Mucho ruido en las muestras
- ▶ Sobre o sub ajuste al entrenar la red neuronal
- ▶ No existen criterios definidos para diseñar la arquitectura de la red neuronal
- ▶ Necesidad de una máquina con buenas prestaciones

Estrategias de solución

- ▶ Aumentar el número de muestras basado en las existentes ya sea estática o dinámicamente
- ▶ Pre-procesar las imágenes
- ▶ Trabajar con estrategias de regularización
 - ▶ L1
 - ▶ L2
- ▶ Poda de neuronas (Crop)
- ▶ Trabajar con redes neuronales artificiales y convolucionales
- ▶ Aprender la generalidad de los defectos y generar nuevas imágenes
- ▶ Trabajar desde varios frentes de pre-procesamiento

Resultados preliminares

Para 29 muestras iniciales: 75% (21) entrenamiento + 25% (8)

```
Epoch 9/10  
42/42 ————— 7s 159ms/step - accuracy: 0.7985 - loss: 0.4786  
Epoch 10/10  
42/42 ————— 7s 156ms/step - accuracy: 0.8353 - loss: 0.4790  
<keras.src.callbacks.history.History at 0x7fa0b1586830>
```

21*20 = 420
muestras de
entrenamiento

```
[ ] 1 #muestras aumentadas para testeo (8 aumentadas a 40)  
2 pred=modeloCNN.predict(X_aumentado)  
3 pred_clases = (pred > 0.5).astype(int)  
4 accuracy = accuracy_score(y_aumentado,pred_clases)  
5 accuracy
```

```
→ 1/1 ————— 9s 9s/step
```

0.7241379310344828

8*10 = 80
muestras de
entrenamiento

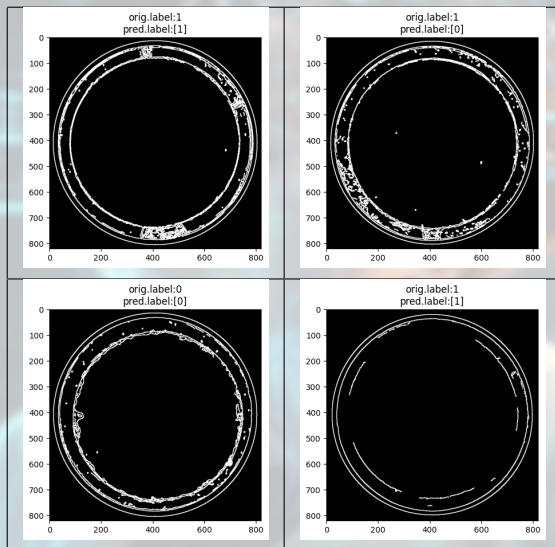
```
[ ] 1 #datos originales que se usan para crear muestras de entrenamiento  
2 X_train_para_test= tf.squeeze(X_train, axis=-1)  
3 pred_train=modeloCNN.predict(X_train_para_test)  
4 pred_clases_train = (pred_train > 0.5).astype(int)  
5 accuracy_train = accuracy_score(y_train, pred_clases_train)  
6 accuracy_train
```

```
→ 1/1 ————— 8s 8s/step
```

0.8095238095238095


21 muestras de
entrenamiento

Resultados preliminares - muestra



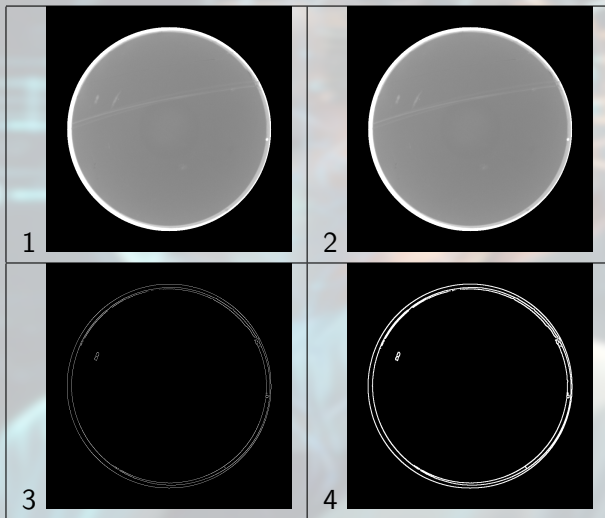
Conclusiones

- ▶ Número de muestras
- ▶ Calidad de las muestras
- ▶ Preprocesamiento de las muestras
- ▶ Heurísticas diseño red
- ▶ Máquina
- ▶ Google Colab



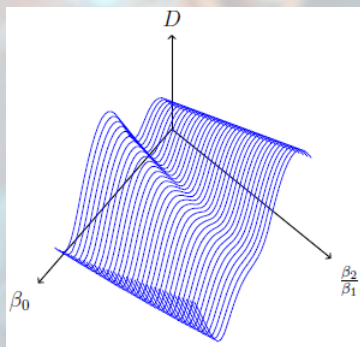
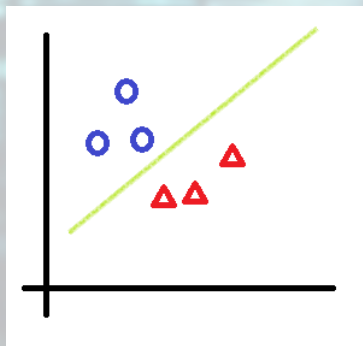
¡Gracias!

Secuencia de pre-procesamiento de las imágenes



Espacio de búsqueda 1

Espacio de búsqueda para un conjunto de datos linealmente separables.



Espacio de búsqueda 2

Conjetura sobre la forma de un espacio de búsqueda con datos en \mathbb{R}^2 no separables.

