

PRESENTACIÓN

PROYECTO MIA

Modelos de Inteligencia Artificial y Sistemas de Aprendizaje Automatico

INTRODUCCIÓN

Model de Descripción de Imágenes Automático

Image Captioning (descripción automática de imágenes) utilizando una arquitectura Encoder-Decoder con las siguientes características:

- **Encoder:** Extrae características visuales usando EfficientNetB3 (pre-entrenado en ImageNet).
- **Decoder:** Genera descripciones usando una LSTM con atención (Bahdanau Attention).
- **Dataset:** Flickr8k

ESQUEMA

```
graph TD; ESQUEMA --> ENCODER; ESQUEMA --> MODELO; ESQUEMA --> DATASET; MODELO --> ARQUIT.; MODELO --> CONFIG.; MODELO --> ENTRE.; ARQUIT. --- CONFIG.; CONFIG. --- ENTRE.
```

ENCODER

MODELO

DATASET

ARQUIT.

CONFIG.

ENTRE.

CONFIG.

DEPENDENCIAS

DATASET

PREPROCESAMIENTO

DATOS

VOCABULARIO
DESCRIPCIONES
CARACTERISTICAS

ARQUIT.

ENCODER

EFFICIENT NET3
LSTM
EMBEDDING

DECODER

LSTM ATENCION
EMBEDDING
MECANISMO DE ATENCIÓN

ENTRE.

HIPERPARAMETROS

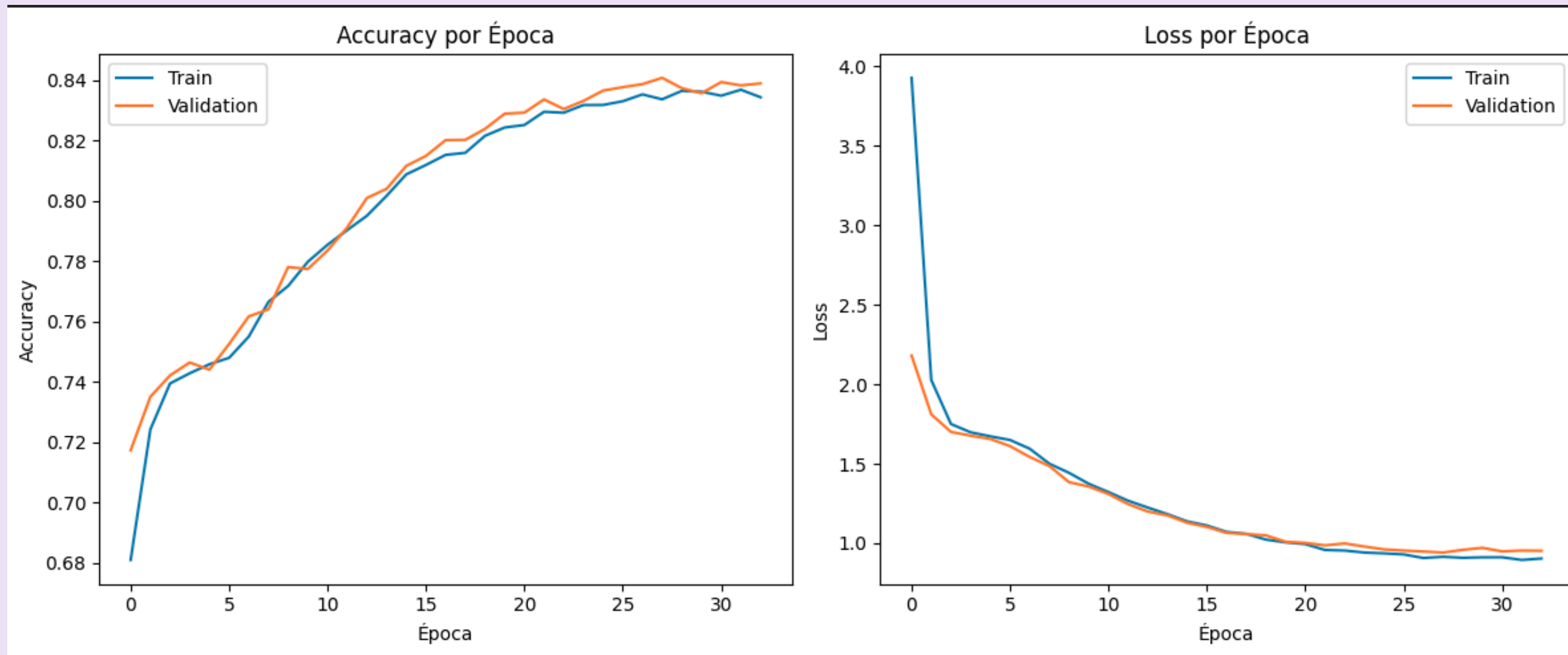
EPOCAS
BATCH SIZE

CALLBACKS

EARLY STOPPING
CHECKPOINT
LEARNING RATE SCHEDULER

EVALUACIÓN

RESULTADO





On grass on field area grass on hind mouth legs side on hind mouth legs side of other



Playing running path beach on field water it

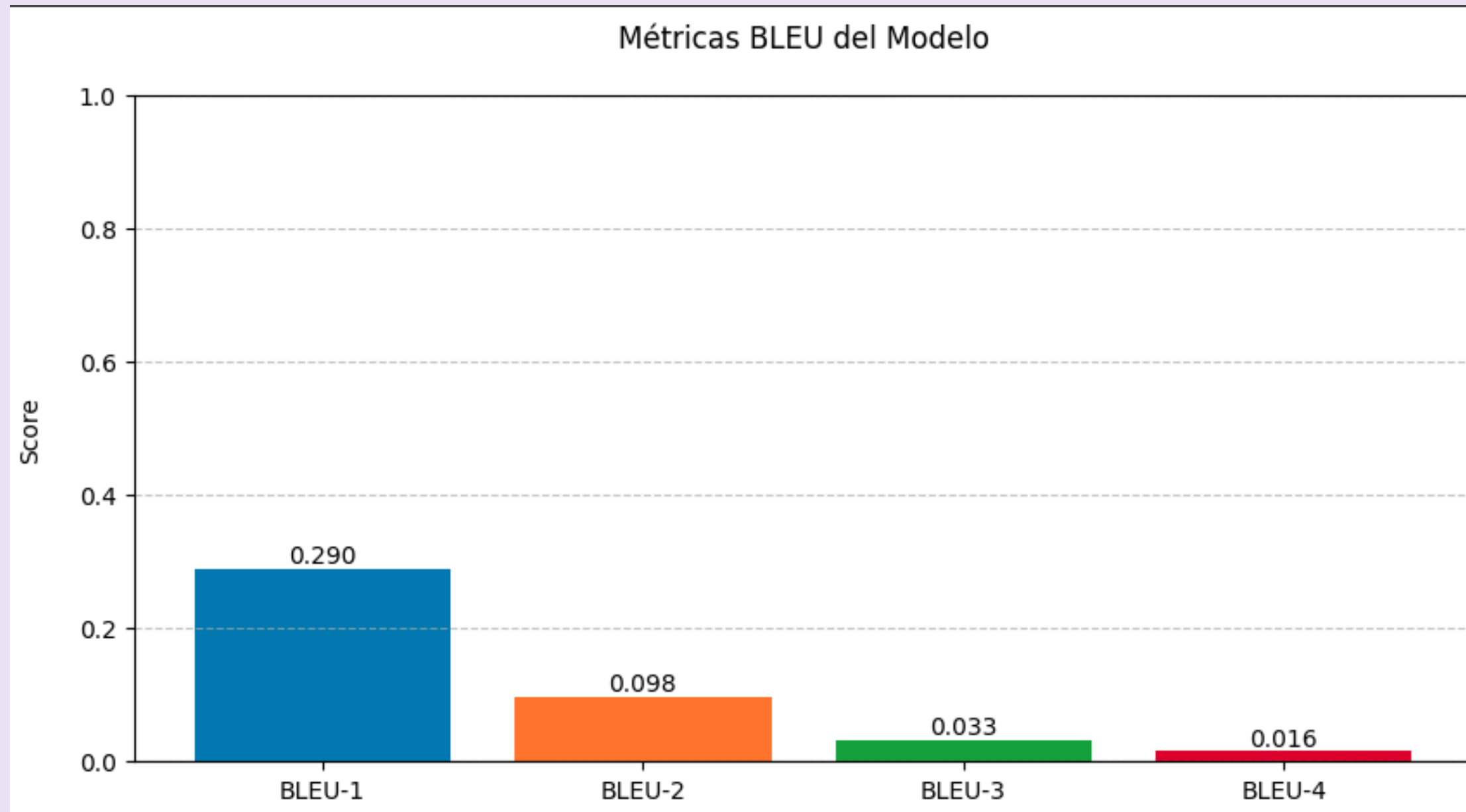


Field and tan of white dogs on water tall snow



On midair on bike legs path on field area of background people him on

RESULTADO



Muchas
GRACIAS