

Use of machine learning algorithms for objects detections in water ways.

Sergio Andrés Sánchez Dueñas
Technological University of
Bolívar.
Computer Science Faculty.
sduenas@utb.edu.co
Cartagena, Colombia.

Andrés Camilo Cardona Arellano.
Technological University of
Bolívar.
Computer Science Faculty.
cardonaa@utb.edu.co
Cartagena, Colombia.

Jesús David Puella Saavedra.
Technological University of
Bolívar.
Computer Science Faculty.
saavedraj@utb.edu.co
Cartagena, Colombia.

Abstract — There are a variety of machine learning algorithms that could be used for objects detections in water ways. However, it is difficult to decide which algorithm will be the best for this task without knowing more about the specific data set that would be used. In general, it is recommended to try a few different algorithms and evaluate which one works best with the data. Therefore, this paper focuses on evaluating the behavior of the variables in order to do a descriptive analysis and standardized the dataset.

I. INTRODUCTION.

Una variable estadística es una función que se puede medir y que es susceptible de ser representada en una tabla o gráfica. Existen dos grandes grupos de variables estadísticas: las continuas y las discretas. Para determinar el nivel de afluencia de una gran cantidad de datos comunes reunidos en un solo conjunto, es importante definir el tipo de variable estadística que maneja el entorno. La variable estadística más utilizada cuando se trata de afluencia de datos es la media. Pero incluso sin conocer el tipo de variable, es importante definir el subconjunto al que hace parte, de este modo podemos catalogarlas dentro de las variables continuas o discretas. Luego de determinar el conjunto al que hace parte, es importante conocer la distribución estadística de los datos que lo forman, ya que muchas de las actividades se basan en función de la varianza, la densidad, y el sistema de correlaciones.

II. DESCRIPTION.

¿Qué tipos de variables hay en el conjunto de datos?

En el análisis estadístico encontramos variables discretas y continuas. Estas se reparten entre todo el conjunto del Dataset. Las variables discretas son las que dan como resultado el algoritmo de reconocimiento, dado que podemos medir el numero

de objetos que se encuentran en el entorno, así como definir el numero de objetos que han sido etiquetados.

Por otra parte, el numero de variables continuas es mayor, esto se debe a que el conjunto de datos para pruebas que abarca la colección hasta el año 2017, no se encuentran etiquetados correctamente, de forma que la cantidad de imágenes preprocesadas evaluadas en este conjunto de datos es completamente incierta. Para dar un breve ejemplo de algunas de las variables continuas a la que debemos enfrentarnos se encuentran:

- El peso (Kb) de las imágenes.
- Cantidad de Imágenes evaluadas por fotogramas.
- Volumen del Agua en mar abierto.
- Entre otras.

¿Cómo son sus distribuciones?

El tipo de distribución estadística que maneja un conjunto de datos preprocesados de objetos comunes es la distribución de probabilidad.

Una distribución de probabilidad es una función que describe el comportamiento de un conjunto de datos aleatorios. La función se puede utilizar para predecir la probabilidad de que un evento ocurra. En este caso, determinar el tipo de objeto que se encuentra en el Fotograma. Las distribuciones de probabilidad se pueden utilizar para el análisis de datos dentro de todo un conjunto relacional de elementos.

¿Siguen faltando valores? ¿Hay variables redundantes?

Siguen faltando valores por etiquetar, es decir, por asignarle un espacio dentro del tipo de objetos que reconoce el dataset COCO. (Common Objects in Context).

Por parte de las variables, no hay imágenes o contenido redundante.

¿Cuáles son las relaciones entre las características?

La relación entre las características depende del tipo de objeto que se está estudiando. Por ejemplo, si se trata de un objeto como una mesa, las características podrían ser la forma, el tamaño, el color, etc. Si se trata de una silla, el ancho, el alto, la distancia, entre otras variables. En el caso de los objetos acuáticos, podríamos definir un faro de luz, una boya, un animal, etc. y cada uno de estos objetos es capaz de tener una característica física evaluable, el peso, el grosor, el alto, entre otras.

¿Observamos valores atípicos?

No, pero aun así se pueden observar datos anómalos o inusuales. Como por ejemplo como esta definido el entorno, que tipo de tiempo favorece la detección del objeto, las sustancias que pueden encontrarse en el mar, entre muchas otras variables.

¿Cómo se correlacionan los diferentes pares de características entre sí? ¿Tienen sentido estas correlaciones?

Las características de los objetos en un dataset están correlacionadas si los valores de una característica tienden a variar de manera similar a los valores de otra característica. Por ejemplo, si dos características están correlacionadas positivamente, esto significa que a medida que los valores de una característica aumentan, los valores de la otra característica también tienden a aumentar. En el proyecto no se evidencian correlaciones entre diferentes pares dado que una imagen no depende de la obtención de otra imagen ya generada.

¿Cuál es la relación entre las características y el objetivo?

Las características permiten que el sistema identifique objetos en el agua, mientras que el objetivo es identificar el objeto en el agua. Las características de detección de objetos en cuerpos de agua se pueden clasificar en cuatro categorías: forma, tamaño, densidad y brillo.

Forma: La forma de un objeto en el agua es una característica importante para la detección de objetos.

Los objetos en el agua pueden tener una forma redonda, ovalada, rectangular o irregular. La forma de un objeto en el agua se puede determinar mediante la medición de la longitud y el ancho del objeto. Tamaño: El tamaño de un objeto en el agua es una característica importante para la detección de objetos. Los objetos en el agua pueden tener un tamaño grande o pequeño. La medición del tamaño de un objeto en el agua se puede realizar mediante la medición de la longitud y el ancho del objeto. Densidad: La densidad de un objeto en el agua es una característica importante para la detección de objetos.

METODOLOGY.

Data, análisis y modelos. Tres características esenciales para la maquetación del sistema de detección de objetos.

Dos características importantes de la detección de objetos son:

1. La capacidad de detectar objetos en movimiento y en tiempo real.
2. La capacidad de detectar objetos de diferentes tamaños y formas.



REFERENCES

- [1] Use of deep learning algorithms for real-time detection of vessels in confined spaces using the Tensorflow framework. S A Sánchez, J Campillo, and J C Martínez-Santos.
- [2] Microsoft COCO: Common Objects in Context
Tsung-Yi Lin¹, Michael Maire², Serge Belongie¹, James Hays³, Pietro Perona²,
Deva Ramanan⁴, Piotr Dollár⁵, and C. Lawrence Zitnick⁵
¹ Cornell ² Caltech ³ Brown ⁴ UC Irvine ⁵ Microsoft Research