

**DETECCIÓN AUTOMÁTICA DEL NIVEL DE ESTRATIFICACIÓN
SOCIOECONÓMICO URBANO USANDO REDES NEURONALES
CONVOLUCIONALES SOBRE IMÁGENES SATELITALES CON
INFORMACIÓN AUMENTADA**

DANIEL ALCIDES CARVAJAL PATIÑO

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOMECHANICAS
ESCUELA DE INGENIERIA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2018

**DETECCIÓN AUTOMÁTICA DEL NIVEL DE ESTRATIFICACIÓN
SOCIOECONÓMICO URBANO USANDO REDES NEURONALES
CONVOLUCIONALES SOBRE IMÁGENES SATELITALES CON
INFORMACIÓN AUMENTADA**

DANIEL ALCIDES CARVAJAL PATIÑO

**TRABAJO DE GRADO PARA OPTAR POR EL TITULO DE
INGENIERIA DE SISTEMAS**

Director

FABIO MARTINEZ CARRILLO

Ph.D. EN INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

Codirector

Ph.D RAUL RAMOS POLLAN

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERIAS FISICOMECHANICAS
ESCUELA DE INGENIERIA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA**

2018

ESPACIO PARA NOTA

ESPACIO PARA CARTA AUTORIZACIÓN USO DE DATOS

CONTENIDO

	pág.
INTRODUCCION	11
1. OBJETIVOS	12
1.1. OBJETIVO GENERAL	12
1.2. OBJETIVOS ESPECIFICOS	12
2. MARCO TEÓRICO	13
2.1. ESTRATIFICACIÓN SOCIAL	13
2.2. MACHINE LEARNING	14
2.2.1. REDES NEURONALES	14
3. DESARROLLO DEL PROYECTO	17
3.1. FUENTES DE DATOS	18
3.2. DATASET	19
3.3. REDES NEURONALES	21
3.4. DETECCION DEL ESTRATO SOCIAL	23
3.5. LA MEJOR CONFIGURACION	26
4. CONCLUSIONES	27
5. RECOMENDACIONES Y TRABAJO FUTURO	28
6. LIMITACIONES Y PROBLEMAS	29

LISTA DE FIGURAS

	pág.
1. Diagrama de red neuronal	15
2. Datos a usar en el proyecto	17
3. Modelo a usar en el proyecto	18
4. Division de Train y Test	20
5. Estrucutra OVERLAECOBO	20
6. Segmentación Semantica	22
7. Ecuacion IoU	24
8. Ejemplo de prediccion	25

LISTA DE TABLAS

	pág.
1. Interfaces	23
2. Interfaces	25

RESUMEN

TITULO: DETECCIÓN AUTOMÁTICA DEL NIVEL DE ESTRATIFICACIÓN SOCIOECONÓMICO URBANO USANDO REDES NEURONALES CONVOLUCIONALES SOBRE IMÁGENES SATELITALES CON INFORMACIÓN AUMENTADA.

AUTORES: DANIEL ALCIDES CARVAJAL PATIÑO.

PALABRAS CLAVE: DANE, Machine learning, Deep Learning, Red Neuronal Convolutacional, .

DESCRIPCION:

La finalidad de este proyecto de grado, es continuar con la investigación que el grupo Conuss ha venido desarrollando durante varios semestres en la creación de una infraestructura nube para la comunidad estudiantil, que permita su uso para el desarrollo de otros proyectos e investigaciones.

Debido a la necesidad de crear diplomados y laboratorios virtuales de computación para fomentar el avance en la formación de ingenieros de calidad, se decide utilizar soluciones de código abierto que permitan el fácil acceso y administración de los servidores físicos y virtualización. Entre las muchas soluciones, se decide optar por OpenStack gracias a su amplia gama de módulos y su abundante comunidad por el cual es respaldado, a su vez, integrando docker como solución para la creación de contenedores.

Todo esto con el fin de que la comunidad estudiantil tenga acceso a recursos de computo que no están al alcance de sus manos, otorgándoles la capacidad de conocer, además de disfrutar, las nuevas tecnologías que hoy por hoy están mejorando y automatizando los procesos de las grandes industrias tecnológicas.

ABSTRACT

TITLE: AUTOMATIC DETECTION OF THE URBAN SOCIOECONOMIC STRATIFICATION LEVEL USING CONVOLUTIONAL NEURAL NETWORKS ON SATELLITE IMAGES WITH INCREASED INFORMATION.

AUTHORS: DANIEL ALCIDES CARVAJAL PATIÑO.

KEYWORDS: Container, Cloud Computing, OpenStack, modules and services.

DESCRIPTION:

The purpose of this thesis is to continue with the research that the Conuss group has incorporated during several semesters in the creation of an infrastructure for the student community that allows its use for the development of other projects and research.

Due to the need to create certified courses and virtual computer labs to promote the advancement in the training of quality engineers, it is decided to use open source solutions that allow easy access and administration of physical servers and virtualization. Among the many solutions, it is decided to opt for OpenStack thanks to its wide range of modules and its abundant community for which it is backed, in turn, integrating docker as a solution for the creation of containers.

All this in order that the student community has access to computing resources that are not available to them, that they can know as well as enjoy the new technologies that are improving today and automating the processes of the large technological industries.

INTRODUCCION

La medición del nivel económico de una zona urbana, actualmente, conlleva un trabajo extenso, como lo expresa el DANE, “en el caso de las revisiones generales urbanas, así como en la estratificación rural se apoya en censos de vivienda”¹. Es decir, se requiere la elaboración de una encuesta de gran tamaño, la cual consume mucho tiempo y personal. Posteriormente, si la encuesta no se realizó usando software de recolección de datos, es necesario realizar su tipeo, lo cual también requiere tiempo. Luego, como lo indica el DANE^{2 3}, el cálculo final del estrato se realiza mediante modelos estadísticos y económicos especialmente calibrados para esta tarea.

En este contexto surgen varias interrogantes respecto a la capacidad de actualización de esta metodología: ¿Qué sucede cuando una ciudad tiene una alta tasa de desarrollo urbano?, ¿Cómo mantiene el gobierno actualizada la información de los estratos ante estas circunstancias?, ¿Que tan efectiva es la metodología actual ante estos casos de alto desarrollo urbano?

Por tanto, el objetivo de este trabajo consiste en seleccionar redes neuronales convolucionales y evaluar su capacidad para determinar automáticamente el estrato socioeconómico usando imágenes satelitales e información adicional (información catastral, presencia y consumo de servicios, etc.), con el fin de presentar una alternativa que haga frente a las inquietudes planteadas.

No es la primera vez que se realiza una predicción del nivel socioeconómico utilizando técnicas de machine learning o Deep learning. Neal Jean en colaboración con varias

¹DANE. Estratificación - Preguntas frecuentes. [en línea]. <https://www.dane.gov.co/files/geoeestadistica/Preguntas_frecuentes_estratificacion.pdf>[citado en 25 de mayo de 2018]

²DANE. Metodología de estratificación. [en línea]. <<http://www.dane.gov.co/index.php/servicios-al-ciudadano/servicios-de-informacion/estratificacion-socioeconomica>>[citado en 25 de mayo de 2018].

³DANE. Procedimiento del cálculo. [en línea]. <<http://www.dane.gov.co/files/geoeestadistica/estratificacion/procedimientoDeCalculo.pdf>>[citado en 25 de mayo de 2018].

personas e instituciones realizó un modelo ^{4 5 6} capaz de predecir el nivel de pobreza en cinco países de África, usando imágenes satelitales y datos extra para dicha tarea. En Colombia, más específicamente en Medellín, también se han realizado modelos ⁷ para determinar niveles socioeconómicos de una zona urbana. En la Universidad EAFIT, usando técnicas de Machine Learning e imágenes satelitales logran medir los índices de pobreza de dicha ciudad.

⁴NEAL jean. Combining satellite imagery and machine learning to predict poverty. [en Línea]. <<http://sustain.stanford.edu/predicting-poverty/>>[citado en 25 de mayo de 2018]

⁵NEAL jean. Combining satellite imagery and machine learning to predict poverty. [en línea]. <<https://github.com/nealjean/predicting-poverty>>[citado en 25 de Mayo de 2018]

⁶NEAL Jean, MARSHALL Burke, † MICHAEL Xie, W. Matthew Davis, DAVID B. Lobell, STEFANO Ermon. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), p. 790-794. 2016

⁷EAFIT. Con imágenes satelitales miden los índices de pobreza en Medellín. [en línea]. <<http://www.eafit.edu.co/investigacion/revistacientifica/edicion-167/Paginas/con-imagenes-satelitales-miden-los-indices-de-pobreza-en-medellin.aspx>>[citado en 25 de mayo de 2018]

1. OBJETIVOS

1.1. OBJETIVO GENERAL

1. Seleccionar y evaluar redes convolucionales para la determinación del nivel socio económico urbano mediante el uso de imágenes satelitales e información adicional.

1.2. OBJETIVOS ESPECIFICOS

1. Identificar fuentes de datos de imágenes satelitales e información adicional.
2. Diseñar y construir datasets integrando los datos obtenidos de las fuentes identificadas.
3. Seleccionar entre distintas arquitecturas de redes neuronales convolucionales existentes en la literatura y repositorios tecnológicos .
4. Entrenar las redes convolucionales probando configuraciones de datasets.
5. Evaluar el desempeño de las redes convolucionales con el uso de los distintos dataset.
6. Elegir la mejor configuración tanto de red convolucional como de conjunto de datos, teniendo en cuenta el desempeño obtenido.

2. MARCO TEÓRICO

2.1. ESTRATIFICACIÓN SOCIAL

“La estratificación social es un fenómeno presente en todas las sociedades. Los miembros se clasifican a sí mismos y a los otros basándose en jerarquías que vienen dadas por diversos factores”⁸ y no es algo nuevo, la antigua Mesopotamia contaba con una división social cuyos miembros iban desde el rey y su familia, en el estrato más alto, hasta los esclavos en el más bajo.

Más que una simple división, la estratificación representa la desigualdad existente en una sociedad. Cada uno de los estratos, niveles o grupos sociales indica la capacidad de acceso a recursos, oportunidades, bienes y servicios por parte de las personas pertenecientes a cada nivel. Dicha estratificación es necesaria para la tarea que llevan a cabo los gobiernos en contra de la desigualdad, dado que se suelen crear programas que benefician a las personas de los niveles más bajos y cobrar mayores impuestos a las personas de los niveles más altos.

En el caso de Colombia, se maneja una estratificación de 6 niveles.

“De éstos 6, los estratos 1, 2 y 3 corresponden a estratos bajos que albergan a los usuarios con menores recursos, los cuales son beneficiarios de subsidios en los servicios públicos domiciliarios; los estratos 5 y 6 corresponden a estratos altos que albergan a los usuarios con mayores recursos económicos, los cuales deben pagar sobrecostos (contribución) sobre el valor de los servicios públicos domiciliarios. El estrato 4 no es beneficiario de subsidios, ni debe pagar sobrecostos, paga exactamente el valor que la empresa defina como costo de prestación del servicio.”⁹

Dicha estratificación no es tarea fácil. Se requiere gran cantidad de variables por cada vivienda, como las características de la zona en la que se ubica, el tamaño, materiales en que fue fabricada, entre otras. Lo que conlleva a una tarea de recolección de datos

⁸GONZÁLEZ Vanessa. ¿Qué es la estratificación social? [en línea]. <<https://www.lifeder.com/estratificacion-social/>>[citado en 31 de mayo del 2018].

⁹DANE, Estratificación - Preguntas frecuentes. Op. cit., p. 1.

bastante amplia.

2.2. MACHINE LEARNING

“Machine learning es una disciplina científica del ámbito de la Inteligencia Artificial que crea sistemas que aprenden automáticamente. Aprender en este contexto quiere decir identificar patrones complejos en millones de datos. La máquina que realmente aprende es un algoritmo que revisa los datos y es capaz de predecir comportamientos futuros. Automáticamente, también en este contexto, implica que estos sistemas se mejoran de forma autónoma con el tiempo, sin intervención humana.”¹⁰

Se han llevado muchos desarrollos y avances en distintos campos usando técnicas de machine learning. Uno de las más recientes y populares fue la máquina de Google AlphaGo que venció al mejor jugador a nivel mundial de GO, Ke Jie. Wikipedia usa técnicas de machine learning para detectar saboteos en su enciclopedia. Otros usos de machine learning consisten en la detección de objetos, patrones o enfermedades incluso predicción de tráfico urbano y precios de bienes.

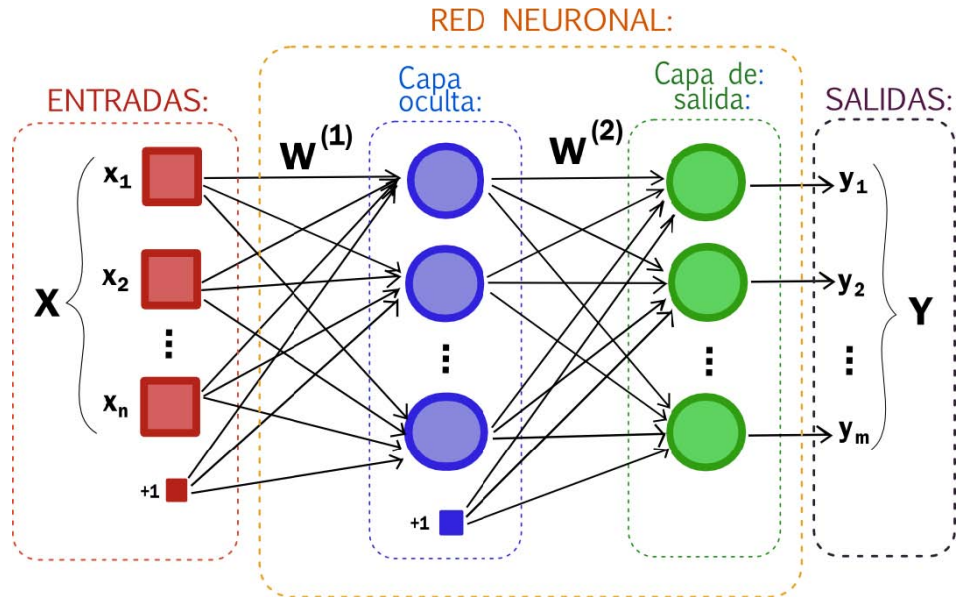
Aunque en el fondo, independientemente del campo en el que se trabaje, las técnicas de machine learning son las mismas, existen tareas mucho más complejas que otras. Por ejemplo, hoy en día, es mucho más fácil predecir el valor de una casa que predecir el valor del dólar. Esta dificultad se debe a que cada tarea requiere “ajustar” los datos y los algoritmos para que se puedan realizar predicciones o detecciones con un nivel de tolerancia aceptable.

2.2.1. redes neuronales es una de las técnicas o algoritmos de machine learning que se pueden emplear en las tareas de predicción o detección. “Una red neuronal es un modelo simplificado que emula el modo en que el cerebro humano procesa la

¹⁰GONZÁLEZ Andrés. ¿Qué es Machine Learning? [en línea]. (Recuperado en 10 oct 2017) <http://cleverdata.io/que-es-machine-learning-big-data/>

información: Funciona simultaneando un número elevado de unidades de procesamiento interconectadas que parecen versiones abstractas de neuronas”¹¹

Figura 1: Diagrama de red neuronal



VÍLCHEZ GARCÍA, Víctor Gabriel. *Estimación y clasificación de daños en materiales utilizando modelos AR y redes neuronales para la evaluación no destructiva con ultrasonidos*. [en línea]. (Recuperado en 24 may 2018) <http://ceres.ugr.es/~alumnos/esclas/>

Una red neuronal está constituida por una serie de capas que se activan con determinadas entradas generando determinadas salidas que podrían ser tomadas o no por otras capas, dependiendo de la “profundidad” de la red. Utilizar redes neuronales consiste en hacer que la misma aprenda examinando las entradas, prediciendo las salidas y haciendo ajustes a los distintos parámetros de la misma, este proceso se repite muchas veces hasta que la red sea capaz de predecir o clasificar con un margen de error tolerable.

¹¹IBM. El modelo de redes neuronales [en línea]. <https://www.ibm.com/support/knowledgecenter/es/SS3RA7_18.0.0/modeler_mainhelp_client_ddita/components/neuralnet/neuralnet_model.html> [citado en 31 de Mayo de 2018]

La idea del proyecto es obtener la configuración de una red neuronal que sea capaz de predecir el nivel socioeconómico de una zona urbana, para esto se planea alimentar la red con imágenes satelitales e información extra de dichas zonas urbanas. Con esta información, ajustes en los parámetros y bastantes iteraciones la red neuronal aprenderá. Los cambios en los parámetros, profundidad y datos usados en la red neuronal son realizados para encontrar la configuración de red neuronal que mejor desempeño presente para la tarea propuesta. Dado que los datos a utilizar son imágenes, es recomendable utilizar “redes neuronales convolucionales” (CNN), las cuales son un tipo de red neuronal que se adapta mejor al uso de imágenes dado que “las CNN eliminan la necesidad de una extracción de características manual, por lo que no es necesario identificar las características utilizadas para clasificar las imágenes. La CNN funciona mediante la extracción de características directamente de las imágenes. Las características relevantes no se entrenan previamente; se aprenden mientras la red se entrena con una colección de imágenes”¹²

¹²MATHWORKS. Aprendizaje profundo [en línea]. <<https://es.mathworks.com/discovery/deep-learning.html>>[citado en 31 de Mayo de 2018]

3. DESARROLLO DEL PROYECTO

Está claro que la recolección de datos mediante encuestas a vivienda es una metodología que conlleva bastante tiempo y que el cálculo del estrato debe tener en cuenta un sin número de variables. Como alternativa se propone usar imágenes satelitales dado que se puede obtener mucha información socioeconómica con el análisis de las mismas. Se pueden identificar patrones, detectar construcciones específicas, clasificar materiales de construcción en los tejados y más, todos estos análisis son variables usadas en el cálculo del estrato.

Como se mencionó existe un par de trabajos sobre el nivel socioeconómico usando técnicas de machine learning. El trabajo de Neal Jean usa redes neuronales convoluciones y Transfer Learning para lograr una predicción de la pobreza en 5 países de África. El trabajo de la EAFIT usa técnicas de machine learning para detectar características que puedan dar a conocer el nivel socioeconómico de una zona. Hay que tener en cuenta que en ambos trabajos la principal fuente de datos son las imágenes satelitales.

A grandes rasgos lo que se planteó realizar con las imágenes satelitales, la información aumentada, la información de los estratos y las redes neuronales, se muestra en las siguientes figuras.

Figura 2: Datos a usar en el proyecto

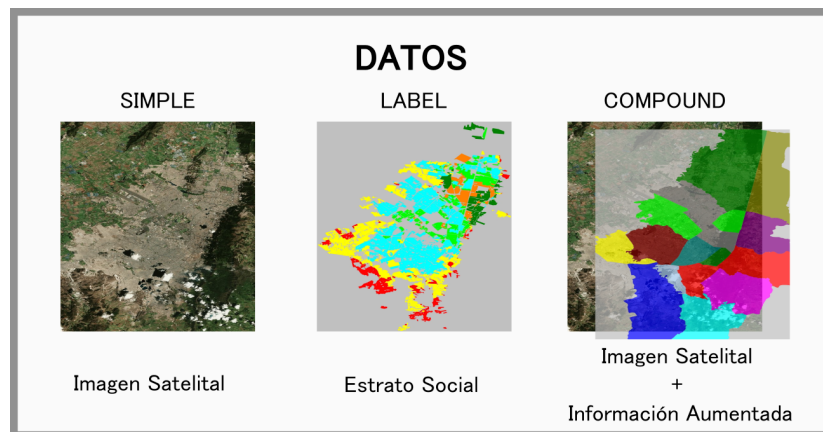
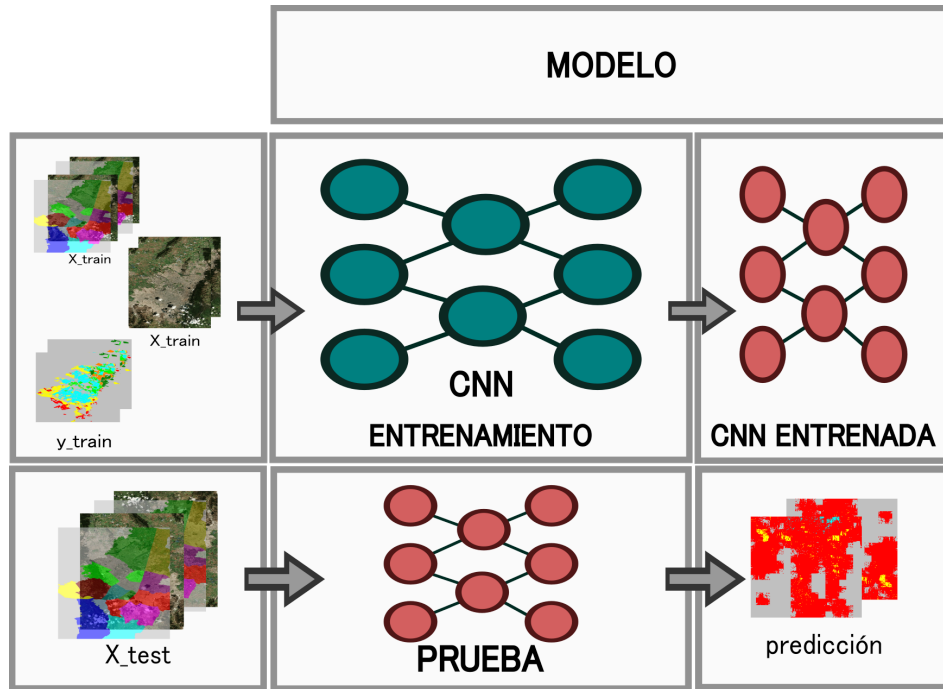


Figura 3: Modelo a usar en el proyecto



El proyecto se desarrolló encontrando fuentes de datos, generando dataset con imágenes satelitales, información aumentada y distintos niveles de zoom, seleccionando redes neuronales y realizando pruebas para determinar la mejor manera de predecir el estrato social. Hay que mencionar que el desarrollo del proyecto está documentado en el siguiente repositorio de GitHub https://github.com/DaielChom/proyecto_uis. El desarrollo se llevó de la siguiente manera.

3.1. FUENTES DE DATOS

Era necesario tener imágenes satelitales, imágenes con los estratos sociales e imágenes con información extra que correspondan geográficamente y que tengan distintos niveles de zoom, como se muestran en la Figura 2. Se usan distintos niveles de zoom para determinar si esto influye en la predicción del estrato social. Existen muchas plataformas para obtener imágenes satelitales tanto de alta como baja resolución, varias de estas

plataformas tienen la opción de insertar información en dichos mapas en forma de marcador, polígono o línea. Esta información por lo general son archivos en formato kml o shp. Gracias a la política de datos abiertos¹³ es posible encontrar archivos kml o shp en distintas páginas web del gobierno. sin embargo no todos los departamentos o ciudades cuentan con la misma cantidad de datos disponibles y menos con datos del estrato social. El único portal web donde se encontró una buena cantidad de información, incluyendo la del estrato social fue en el portal de mapas de Bogotá, disponible en <http://mapas.bogota.gov.co>.

No fue posible descargar los archivos kml mostrados en la plataforma, por ende, y aprovechando la licencia Creative Commons que presentan los datos, se crearon nuevos archivos kml usando los mapas de la plataforma como guía. Los archivos kml realizados están disponibles en <https://drive.google.com/open?id=15VnvN6ZRTbsqqd9k13ukNWBj3oqADsy0>. se crearon dos kml, uno con información de los estratos sociales y otro con los índices de condiciones de seguridad nocturna por localidad que tienen las mujeres, mas específicamente los datos de la categoría *riesgo alto* de dicho mapa.

3.2. DATASET

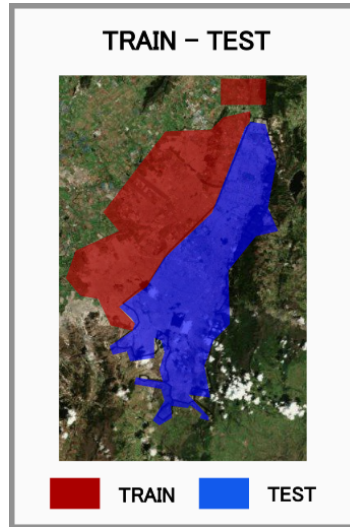
Como en la mayoría de trabajos de machine learning se debe tener un dataset, un conjunto de datos organizados y con una estructura dividida en train (datos para entrenamiento) y test (datos para pruebas) con los cuales se entrena o prueba un algoritmo de aprendizaje. El dataset que se construyó lleva por nombre OVERLAECOBO y cuenta con 3 tipos imágenes (Figura 2) geográficamente correspondidas y de distintos niveles de zoom o acercamiento satelital. Las imágenes del dataset fueron obtenidas usando el software Ruso SASplanet¹⁴, los archivos kml diseñados y *Bing maps* como fuente de imágenes satelitales, fuente ofrecida en SASplanet. Se tomo el mapa de Bogotá y se trazó una línea que dividiera la ciudad en dos partes, de tal manera que en ambas hubiera información de todos los estratos, una parte se escogió para train y otra la para

¹³GOBIERNO DIGITAL. Datos Abiertos [en línea] <<http://estrategia.gobiernoenlinea.gov.co/623/w3-article-9407.html>>[citado en 30 de Mayo de 2018]

¹⁴Disponible en <http://www.sasgis.org/>

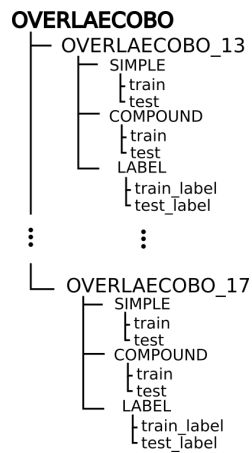
test, como se muestra en la siguiente imagen.

Figura 4: Division de Train y Test



Se extrajeron imágenes de 7 Zooms diferentes, del 13 al 20, cada nivel de zoom cuenta con imágenes satelitales, de información extra (Riesgo nocturno para las mujeres) y de estrato social. Cada nivel está dividido en train y test. OVERLAECOBO cuenta con la siguiente estructura.

Figura 5: Estructura OVERLAECOBO

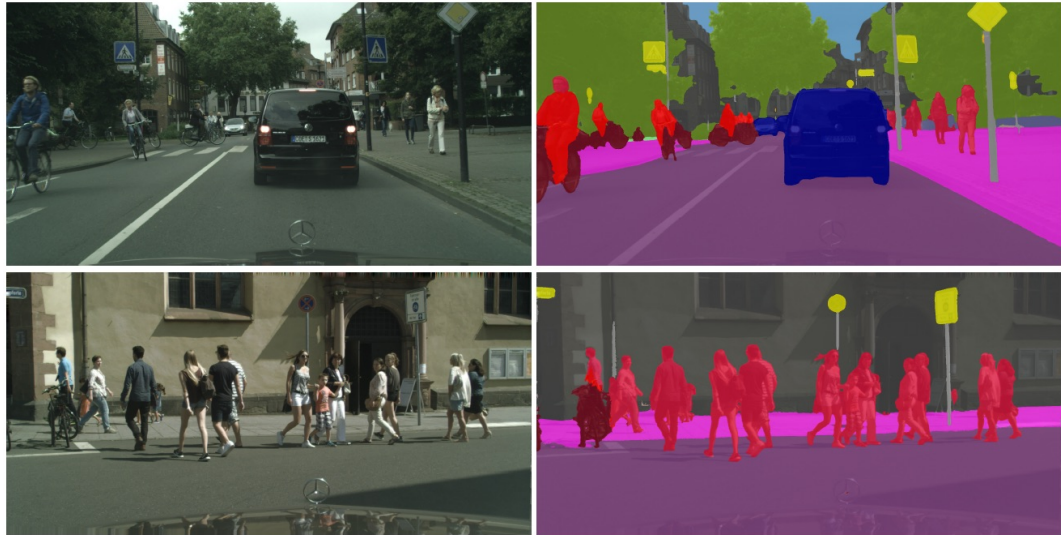


Una vez el dataset estuvo elaborado se prosiguió con la búsqueda de redes neuronales y posteriormente pruebas para determinar la mejor configuración para la detección del estrato social.

3.3. REDES NEURONALES

Hay que mencionar que la metodología que se usó para la detección del estrato social se conoce en machine learning como Segmentación Semántica (Figura 6), la cual consiste en segmentar una imagen con el fin de delimitar un objeto o diferenciar entre elementos. En Segmentación semántica, en el proceso de entrenamiento se tienen dos imágenes, la imagen que se quiere segmentar (X_{train}) y la imagen segmentada (y_{train}) y con estas se realiza el aprendizaje. En el proceso de pruebas con imágenes que se requieren segmentar (X_{test}) se crea una nueva imagen segmentada, luego en el proceso de medición del error con las imágenes segmentadas creadas se comparan con las imágenes segmentadas de dataset (y_{test}). En otras palabras, cada pixel RGB de una imagen tiene su propia etiqueta. La segmentación semántica es una técnica muy utilizada en el campo de la conducción autónoma, ya que representa una buena metodología para descubrir los distintos objetos que percibe la cámara del vehículo y así tomar decisiones.

Figura 6: Segmentación Semantica



ABHIJIT Kundu, VIBHAV Vineet, and VLADLEN Koltun Feature Space Optimization for Semantic Video Segmentation [en línea] <<http://vladlen.info/publications/feature-space-optimization-for-semantic-video-segmentation/>> [citado en 01 de junio de 2018]

Existe muchas redes neuronales convolucionales con las que se puede aplicar Segmentación semántica, entre ellas SegNet, uNet y AlexNet. En GitHub hay varios repositorios con implementaciones de estas redes; como es de esperarse estos repositorios, aunque implementen la misma red neuronal, tienen varias diferencias; implementados en distintos lenguajes, usan diferentes librerías e incluso entrenan con diferentes dataset, sin embargo, cumplen la tarea de Segmentación semántica que proponen. Tras una búsqueda y selección entre varios repositorios para seleccionar la mejor red neuronal o la mejor implementación se encontró un Framework que permite usar distintas redes neuronales para la misma tarea. Dicho framework está disponible en el siguiente enlace <https://github.com/GeorgeSeif/Semantic-Segmentation-Suite> y es el que se usó para las pruebas.

3.4. DETECCION DEL ESTRATO SOCIAL

Se utilizaron 3 niveles de ZOOM para las pruebas y de cada nivel 2 configuraciones, *SIMPLE* y *COMPOUND* (satelital, satelital + informacion aumentada) y para cada configuracion se realizaron 11 pruebas variando la red neuronal y los parametros *epoch* y *batch size*. En total se realizaron 66 pruebas.

Tabla 1: Pruebas realizadas por cada configuracion de dataset


Pruebas			
Nº	RED 1	EPOCH	BATCH SIZE
1	FC-DenseNet56	50	1
2	FC-DenseNet56	70	1
3	Encoder-Decoder	70	1
4	Encoder-Decoder	70	2
5	Encoder-Decoder-Skip	70	1
6	Encoder-Decoder-Skip	70	2
7	FRRN-A	70	1
8	MobileUNet	70	1
9	MobileUNet	70	2
10	MobileUNet-Skip	70	1
11	MobileUNet-Skip	70	2

una vez realizado el entrenamientos de cada red neuronal se midie la capacidad de cada una para peredeccir el estrato social (generar las imagenes *LABEL*) utilizando el indice de jaccard, mas conocido como *IoU - Interseccion over Union*, una metrica¹⁴ muy utilizada en la deteccion de objetos y la segmentacion. El indice de jaccard consiste en encontrar el grado de similitud entre dos conjuntos, en este caso entre dos imagenes, calculando el cociente entre la interseccion y union de los dos conjuntos. En el caso de

¹⁴ROSEBROCK Adrian. Intersection over Union (IoU) for object detection [en linea] <<https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>>[citado en 17 de Julio de 2018]

segmentacion semantica se usan las areas cubierta por cada clase en las iamgenes para calcular dicho coeficiente, esto se explica mejor en la siguiente imagen.

Figura 7: Ecuacion IoU

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


ROSEBROCK Adrian. Intersection over Union (IoU) for object detection [en linea] <<https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/> [citado en 17 de julio de 2018]

Cada red entrenada genero una cantidad de imagenes equivalente a las de *test* del dataset con el que se entreno. Con estas imagenes generadas y las imagenes del dataset se calculo el indice de jaccard para cada una de las 66 pruebas. Como se mencionó dichos procesos se pueden observar en el repositorio del proyecto. https://github.com/DaielChom/proyecto_uis/

Figura 8: Ejemplo de prediccion

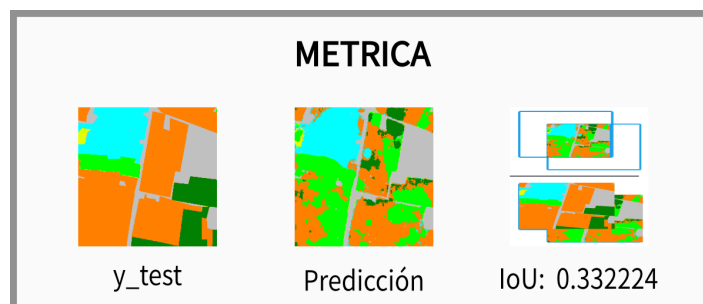


Tabla 2: Resultados de cada entrenamiento

Pruebas								
Prueba	Duración (s)	IoU	Prueba	Duración (s)	IoU	Prueba	Duración (s)	IoU
1	1696	-	23	3029	-	45	6768	-
2	1860	-	24	3510	-	46	9545	-
3	1910	-	25	3468	-	47	9305	-
4	1705	-	26	3038	-	48	9659	-
5	2480	-	27	4523	-	49	11346	-
6	1984	-	28	3229	-	50	7780	-
7	2591	-	29	4272	-	51	11429	-
8	1083	-	30	2267	-	52	6996	-
9	1137	-	31	2016	-	53	6559	-
10	934	-	32	2209	-	54	7733	-
11	917	-	33	2142	-	55	6629	-
12	1345	-	34	2524	-	56	7151	-
13	1993	-	35	3510	-	57	9542	-
14	1962	-	36	4518	-	58	9672	-
15	2052	-	37	3231	-	59	7877	-
16	2088	-	38	3467	-	60	9527	-
17	1706	-	39	3110	-	61	8212	-
18	1677	-	40	3793	-	62	11611	-
19	900	-	41	2914	-	63	7125	-
20	810	-	42	2216	-	64	6547	-
21	1357	-	43	2256	-	65	7763	-
22	876	-	44	2160	-	66	7119	-

3.5. LA MEJOR CONFIGURACION

asd

4. CONCLUSIONES

5. RECOMENDACIONES Y TRABAJO FUTURO

6. LIMITACIONES Y PROBLEMAS

BIBLIOGRAFIA

DANE. Estratificación - Preguntas frecuentes. [en línea]. <https://www.dane.gov.co/files/geoestadistica/Preguntas_frecuentes_estratificacion.pdf>[citado en 25 de Mayo de 2018]

DANE. Metodología de estratificación. [en línea]. <<http://www.dane.gov.co/index.php/servicios-al-ciudadano/servicios-de-informacion/estratificacion-socioeconomica>>[citado en 25 de Mayo de 2018]

DANE. Procedimiento del cálculo. [en línea]. <<http://www.dane.gov.co/files/geoestadistica/estratificacion/procedimientoDeCalculo.pdf>>. [citado en 25 de Mayo de 2018].

NEAL jean. Combining satellite imagery and machine learning to predict poverty. [en Linea]. <<http://sustain.stanford.edu/predicting-poverty/>>[citado en 25 de Mayo de 2018].

NEAL jean. Combining satellite imagery and machine learning to predict poverty. [en línea]. <<https://github.com/nealjean/predicting-poverty>>[citado en 25 de Mayo de 2018]

NEAL Jean, MARSHALL Burke, † MICHAEL Xie, W. Matthew Davis, DAVID B. Lobell, STEFANO Ermon. Combining satellite imagery and machine learning to predict poverty. Science 353 (6301), p. 790-794. 2016

EAFIT. Con imágenes satelitales miden los índices de pobreza en Medellín. [en línea]. <<http://www.eafit.edu.co/investigacion/revistacientifica/edicion-167/Paginas/con-imagenes-satelitales-miden-los-indices-de-pobreza-en-medellin.aspx>>[citado en 25 de Mayo de 2018]

GONZÁLEZ Vanessa. ¿Qué es la estratificación social? [en línea]. <<https://www.lifeder.com/estratificacion-social/>>[citado en 31 de Mayo del 2018].

IBM. El modelo de redes neuronales [en línea]. <https://www.ibm.com/support/knowledgecenter/es/SS3RA7_18.0.0/modeler_mainhelp_client_ddita/components/neuralnet/neuralnet_model.html> [citado en 31 de Mayo de 2018]

MATHWORKS. Aprendizaje profundo [en línea]. <<https://es.mathworks.com/discovery/deep-learning.html>> [citado en 31 de Mayo de 2018]

GOBIERNO DIGITAL. Datos Abiertos [en línea] <<http://estrategia.gobiernoenlinea.gov.co/623/w3-article-9407.html>> [citado en 30 de Mayo de 2018]