

Tabla de Contenido

1. Introducción	1
1.1. Contexto y motivación	1
1.2. Problema a Resolver	2
1.3. Solución propuesta	2
1.4. Objetivos	3
1.5. Metodología	3
2. Marco Teórico	4
2.1. Redes Neuronales Artificiales	4
2.1.1. Perceptrón	4
2.1.2. Convolución	5
2.1.3. Redes Neuronales Convolucionales	5
2.1.4. ResNet-152	6
2.2. Descriptores Visuales	7
2.2.1. Deep Feature	7
2.3. Algoritmos de Clusterización	7
2.3.1. Density-Based Spatial Clustering of Applications with Noise (DBSCAN)	8
2.3.2. Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN)	9
2.4. Métricas de evaluación	11
2.4.1. Mutual Information (MI)	11
2.4.2. Adjusted Mutual Information (AMI)	12

2.4.3.	Rand Index (RI)	13
2.4.4.	Adjusted Rand Index (ARI)	13
2.4.5.	Precision y Recall	14
2.4.6.	Precision at k (P@k)	14
2.4.7.	Intersection Over Union (IOU)	15
2.5.	Distancia de Levenshtein	15
2.6.	Term Frequency - Inverse Document Frequency (TF-IDF)	15
2.7.	Reducción de Dimensionalidad	16
2.7.1.	Principal Component Analysis (PCA)	16
2.7.2.	t-distributed stochastic neighbor embedding (t-SNE)	17
2.7.3.	Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP)	18
3.	Recopilación de Antecedentes	19
3.1.	Estado del arte	19
3.1.1.	Clasificación	19
3.1.2.	Deep Features	21
3.1.3.	Visualización de espacios de alta dimensión	21
4.	Descripción e Implementación de la Solución	22
4.1.	Descripción General de la Solución	22
4.2.	Implementación	22
4.2.1.	Conjuntos de datos	22
4.2.2.	Etiquetado de búsquedas y detección de Zero-Hits	23
4.2.3.	Clusterización de etiquetas	26
4.2.4.	Clusterización de descriptores	26
4.2.5.	Selección de clústeres	27
5.	Experimentos y Análisis de Resultados	29

5.1. Diseño de Experimentos y Evaluación	29
5.1.1. Datos para la evaluación	29
5.1.2. Experimentos	30
5.2. Resultados y Análisis	32
5.2.1. Detección de Zero-Hits	32
5.2.2. Etiqueta de búsquedas visuales	33
5.2.3. Clusterización	33
5.2.4. Selección de clústeres	34
6. Conclusiones y Trabajo Futuro	38
Bibliografía	41
Anexo A	42
Anexo B	44