

## 1. Pengenalan Zero-Shot Learning (ZSL)

### Apa itu Generalisasi?

- **Generalisasi** adalah kemampuan model untuk mengenali data yang belum pernah dilihat sebelumnya, seperti mengenali jenis kucing yang baru meskipun hanya mempelajari karakteristik umum tentang "kucing" selama pelatihan.
- **Domain Adaptation (DA)**: Teknik untuk menyesuaikan model agar dapat bekerja pada data baru yang berasal dari domain yang berbeda, seperti mengenali kucing kartun setelah pelatihan pada gambar kucing nyata.
- **Zero-Shot Learning (ZSL)**: Pendekatan yang memungkinkan model untuk mengenali objek baru yang tidak pernah dilihat selama pelatihan, misalnya mengenali hewan yang tidak pernah dilatih seperti llama atau elang.

### Definisi Zero-Shot Learning (ZSL)

- ZSL adalah pendekatan di mana model diuji dengan data dari kelas yang tidak pernah dilihat selama pelatihan.
- **Generalized Zero-Shot Learning (GZSL)**: Pengujian dengan kombinasi kelas yang sudah dilihat dan yang belum dilihat. Pendekatan ini lebih mencerminkan situasi dunia nyata di mana model harus menangani data baru yang tidak dikenali sebelumnya.

## 2. Sejarah Zero-Shot Learning

- **2008**: Konsep awal ZSL muncul pada konferensi AAAI.
- **2009**: Geoffrey Hinton memperkenalkan istilah Zero-Shot Learning.
- **2021**: OpenAI meluncurkan **CLIP**, yang merevolusi cara kerja ZSL dalam visi komputer dengan memanfaatkan hubungan antara gambar dan teks.

## 3. Bagaimana Zero-Shot Learning Bekerja?

Untuk mengenali objek baru, model memerlukan **informasi tambahan** berupa:

- **Vektor Atribut**: Fitur seperti "bergaris" atau "berbulu".
- **Deskripsi Teks**: Penjelasan teks terkait gambar yang dapat memberikan konteks.
- **Vektor Label Kelas**: Representasi embedding dari label kelas objek yang dikenal.

Proses pelatihan ZSL dilakukan dengan memetakan fitur gambar ke fitur semantik, kemudian saat pengujian, model akan mencocokkan gambar dengan label yang paling dekat di ruang semantik, biasanya menggunakan teknik seperti **k-nearest neighbor** (k-NN).

## 4. Perbedaan ZSL dengan Transfer Learning

- **ZSL** adalah bentuk transfer learning, tetapi diterapkan pada **kelas yang belum pernah dilihat**. Ini berbeda dengan transfer learning tradisional yang bekerja pada domain target yang masih berhubungan dengan data pelatihan.

## 5. Metode ZSL dalam Visi Komputer

### 1. Embedding-based Methods:

- Model ini memproyeksikan gambar dan fitur semantik ke dalam ruang embedding yang sama. Contoh: **CLIP**, yang menghubungkan teks dan gambar untuk meningkatkan klasifikasi.

### 2. Generative-based Methods:

- Menggunakan model generatif, seperti **CVAE** (Conditional Variational Autoencoder), untuk menghasilkan gambar sintetis dari kelas yang tidak pernah dilihat, kemudian melatih model untuk mengenali gambar-gambar ini.

## 6. ZSL vs Pendekatan Lain

- **Domain Adaptation (DA)**: Mengatasi perbedaan antar domain (misalnya, gambar nyata vs. kartun), tetapi membutuhkan data berlabel.
- **Open Set Recognition (OSR)**: Mendeteksi apakah data adalah bagian dari kelas yang dilatih atau tidak, namun tidak mengenali kelas baru.
- **Out-of-Distribution (OOD) Detection**: Mendeteksi data yang tidak sesuai dengan distribusi data pelatihan.
- **Open Vocabulary Learning (OVL)**: Variasi lanjutan ZSL yang dapat menangani kelas tak terbatas dan data baru.

## 7. ZSL vs Generalized Zero-Shot Learning (GZSL)

- **ZSL**: Model hanya diuji pada kelas yang tidak pernah dilihat selama pelatihan.
- **GZSL**: Model diuji dengan gabungan kelas yang pernah dilihat dan yang tidak pernah dilihat. Lebih sulit karena model harus memisahkan data yang berasal dari kelas baru atau lama.

## 8. Inductive vs Transductive ZSL

- **Inductive ZSL**:
  - Model hanya dilatih dengan data dari kelas yang telah dilihat.
  - Fokus pada generalisasi pola untuk mengklasifikasikan kelas yang belum dilihat.
- **Transductive ZSL**:
  - Model memiliki akses ke beberapa informasi tambahan tentang kelas yang belum dilihat selama pelatihan.
  - Menggunakan informasi tersebut untuk melatih model lebih efektif dalam menggeneralisasi data kelas yang tidak dilihat.

## 9. Semantic Embeddings dalam ZSL

- **Apa itu Semantic Embeddings?**

- **Semantic embeddings** adalah representasi vektor dari informasi semantik yang menggambarkan makna dan hubungan kontekstual data dalam ruang berdimensi tinggi.
- Proses embedding melibatkan pemetaan informasi semantik ke dalam bentuk vektor angka menggunakan model seperti **Word2Vec** atau **GloVe**.

- **Penggunaan Semantic Embeddings dalam ZSL:**

- **Pelatihan:** Model belajar untuk menghubungkan fitur visual gambar dengan embedding semantik mereka untuk meminimalkan jarak antara keduanya.
- **Inferensi:** Model memproyeksikan gambar dari kelas yang belum dilihat ke ruang semantik dan mencari kelas dengan embedding terdekat.

## 10. Perbandingan ZSL dengan CLIP

- **Kesamaan:**

- Baik ZSL maupun CLIP memiliki tujuan yang sama, yaitu mengenali gambar dari kelas yang tidak dikenal selama pelatihan.
- Keduanya mengandalkan **informasi tambahan** untuk meningkatkan klasifikasi.

- **Perbedaan:**

- **ZSL tradisional** bergantung pada **embedding semantik** yang telah ditentukan sebelumnya.
- **CLIP** mempelajari hubungan langsung antara **teks dan gambar** secara bersamaan, memungkinkan generalisasi yang lebih luas tanpa embedding khusus untuk setiap tugas.

## 11. Dataset Evaluasi ZSL

Beberapa dataset yang umum digunakan untuk evaluasi metode ZSL adalah:

1. **Animal with Attributes (AWA):** 30.475 gambar, 50 kelas hewan, atribut berbasis transfer.
2. **Caltech-UCSD-Birds (CUB):** 11.788 gambar, 200 subkategori burung, atribut biner dan deskripsi teks.
3. **Sun Database (SUN):** 130.519 gambar, 899 kategori pengenalan adegan.
4. **Attribute Pascal and Yahoo (aPY):** 15.339 gambar, 32 subkategori (hewan, objek, kendaraan).
5. **ILSVRC (ImageNet):** Dataset besar untuk klasifikasi gambar dan deteksi objek.

## 12. Kesimpulan

Zero-Shot Learning (ZSL) adalah paradigma yang memungkinkan model untuk mengenali kelas baru yang tidak dilihat sebelumnya dengan menggunakan informasi semantik tambahan. Meskipun ada beberapa tantangan seperti generalisasi dan akurasi pada kelas yang tidak

dilihat, perkembangan terbaru seperti CLIP dan pendekatan berbasis generative memberikan potensi besar untuk memperluas kemampuan model dalam mengatasi masalah ini. Dengan terus menggunakan dan mengembangkan ZSL, kita dapat meningkatkan kemampuan model untuk menghadapi dunia nyata yang penuh dengan data baru dan tidak terstruktur.