Stamp Verification Using Fine Tune Segformer Transformer

Group Members:

Ali Shah Naushad	**20k-1078**
Yahya Naeem	**20k-1065**
Mehma Waseem	**20k-0261**

Submitted To:
Miss Sumaiyah Zahid

Objective:

The objective of this project is to develop an image segmentation model using a pre-trained model from Hugging Face on the StaVer dataset. This model will accurately segment images of stamps and distinguish between images containing a stamp and those without a stamp.

Image segmentation is a fundamental task in computer vision, where the goal is to divide an image into multiple regions or segments, with each segment representing a different object or region in the image. In our case, we are focusing on segmenting images of stamps, distinguishing them from images without stamps.

To achieve this, we plan to utilize a pre-trained Segformer transformer model from Hugging Face, a well-known platform for natural language processing and computer vision models. By fine-tuning this pre-trained model on the StaVer dataset, we aim to enhance its ability to accurately segment images and classify them based on whether they contain a stamp or not.

Problem Statement:

The problem we are addressing is the need to automate the process of identifying images containing stamps in various applications such as document processing and mail sorting. Currently, this task often requires manual inspection, which is time-consuming and labor-intensive.

By developing an image segmentation model, we aim to create a system that can automatically identify and segment images containing stamps. This will enable efficient verification of documents and sorting of mail, saving time and effort for organizations that deal with large volumes of documents and mail.

The StaVer dataset, available on Kaggle, provides a valuable resource for training our model. It consists of images of stamps along with their corresponding labels indicating whether they contain a stamp or not. By training our model on this dataset, we can teach it to recognize the features of stamps and accurately segment them from other parts of the image.

Motivation:

The proposed image segmentation fine tuned segformer model has several potential applications in various domains. In document processing, the model can be used to automatically identify whether a scanned document contains a stamp, which is useful for verifying the authenticity of the document. This can be particularly important for legal documents, contracts, and other sensitive materials.

Similarly, in mail sorting, the model can be employed to automatically identify envelopes with stamps, enabling automated processing of mail. This can significantly speed up the sorting

process in postal services and reduce errors associated with manual sorting.

Furthermore, the development of an accurate image segmentation model contributes to the field of computer vision and image processing. By automating the process of identifying images containing stamps, we can save time and effort compared to manual inspection. Moreover, the results of this project can be extended to other similar domains, such as identifying images containing specific objects or features.

Overall, the proposed project has the potential to make a significant contribution to the field of computer vision and image segmentation. By leveraging pre-trained models from Hugging Face and fine-tuning them on the StaVer dataset, we aim to achieve state-of-the-art results in stamp image segmentation. Additionally, the insights gained from this project can be extended to other image segmentation tasks in the future, further advancing the capabilities of computer vision systems.

Methodology:

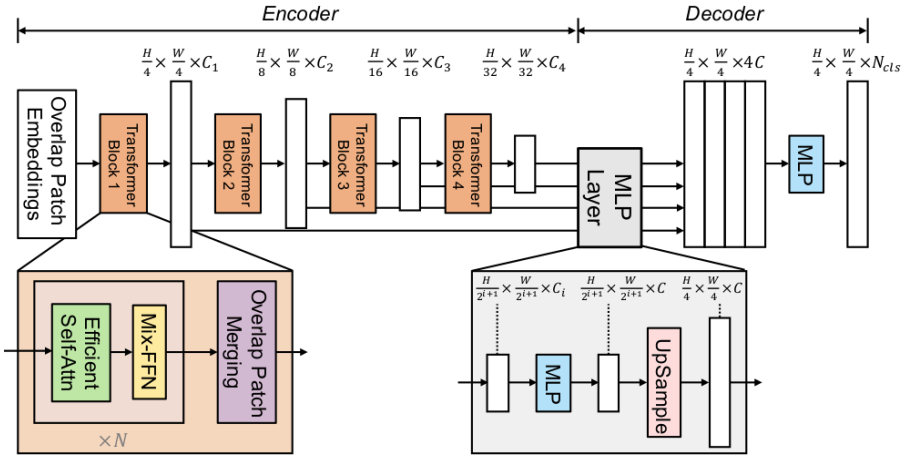

Fig: Segfromer Architecture

1. **Data Acquisition and Preparation:**
    a. We begin by obtaining the StaVer dataset from Kaggle, which contains images of stamps along with corresponding labels indicating whether they contain a stamp or not.
    b. Next, we leverage the segment.ai platform to manually create segmentation masks around stamps in the images. This step is crucial for training our model as it provides precise annotations for the segmentation task.
    c. After annotating the images, we organize the dataset into appropriate directories, separating images and their corresponding segmentation masks, ready for model training.

2. **Dataset Publication on Hugging Face:**
    a. We then proceed to upload the annotated dataset to our Hugging Face profile. This involves creating a release of the dataset on the platform, providing

necessary metadata such as description, license, and tags for easy discovery.

3. **Model Selection and Pre-training:**
    a. For the image segmentation task, we opt for the Segformer transformer model, available on Hugging Face's model hub. Specifically, we choose the Segformer B0 variant for its balance between performance and computational requirements.
    b. We initialize the Segformer B0 model with its pre-trained weights, which have been trained on large-scale image datasets, enabling it to capture general features relevant to image segmentation tasks.
4. **Fine-tuning on StaVer Dataset:**
    a. Using the annotated StaVer dataset, we fine-tune the pre-trained Segformer B0 model on the task of stamp image segmentation.
    b. During fine-tuning, we adjust the model's parameters to better adapt to the characteristics of stamp images, emphasizing features relevant to stamp identification and segmentation.
    c. The fine-tuning process involves optimizing the model's weights through backpropagation while minimizing a suitable loss function, such as binary cross-entropy, which compares the predicted segmentation masks with the ground truth annotations.
5. **Publication of Fine-tuned Model:**
    a. Once fine-tuning is complete, we upload the fine-tuned Segformer B0 model along with its performance metrics to our Hugging Face profile.
    b. Performance metrics may include evaluation results on validation data, such as accuracy, precision, recall, and IoU (Intersection over Union), providing insights into the model's effectiveness in segmenting stamps from images.
6. **Inference and Evaluation:**
    a. Finally, we load the fine-tuned Segformer B0 model for inference on new images containing stamps.
    b. During inference, the model generates segmentation masks for input images, delineating regions corresponding to stamps.
    c. We evaluate the model's performance on a separate test set, measuring metrics such as accuracy, IoU, and other score to assess its ability to accurately segment stamps.

Results:

Training Loss	Epoch	Step	Validation Loss	Mean Iou	Mean Accuracy	Overall Accuracy	Accuracy Unlabeled	Accuracy Stamp	Iou Unlabeled
0.4566	0.8333	20	0.4738	0.1430	0.2860	0.2860	nan	0.2860	0.0
0.3076	1.6667	40	0.3046	0.1307	0.2614	0.2614	nan	0.2614	0.0
0.2373	2.5	60	0.2226	0.0604	0.1209	0.1209	nan	0.1209	0.0
0.2184	3.3333	80	0.2220	0.1942	0.3884	0.3884	nan	0.3884	0.0
0.1578	4.1667	100	0.1704	0.2468	0.4936	0.4936	nan	0.4936	0.0
0.1412	5.0	120	0.1269	0.0376	0.0751	0.0751	nan	0.0751	0.0
0.1109	5.8333	140	0.1076	0.2741	0.5483	0.5483	nan	0.5483	0.0
0.106	6.6667	160	0.0892	0.0583	0.1166	0.1166	nan	0.1166	0.0
0.0899	7.5	180	0.0747	0.0173	0.0346	0.0346	nan	0.0346	0.0
0.0794	8.3333	200	0.0683	0.0189	0.0378	0.0378	nan	0.0378	0.0
0.0741	9.1667	220	0.0639	0.0981	0.1963	0.1963	nan	0.1963	0.0
0.0832	10.0	240	0.0559	0.0599	0.1198	0.1198	nan	0.1198	0.0
0.0575	10.8333	260	0.0527	0.0769	0.1538	0.1538	nan	0.1538	0.0
0.05	11.6667	280	0.0502	0.0852	0.1704	0.1704	nan	0.1704	0.0
0.0523	12.5	300	0.0446	0.1038	0.2076	0.2076	nan	0.2076	0.0
0.0481	13.3333	320	0.0431	0.0956	0.1913	0.1913	nan	0.1913	0.0
0.0471	14.1667	340	0.0420	0.1330	0.2660	0.2660	nan	0.2660	0.0
0.042	15.0	360	0.0412	0.1124	0.2248	0.2248	nan	0.2248	0.0

- Loss: 0.0365
- Mean Iou: 0.1372
- Mean Accuracy: 0.2744
- Overall Accuracy: 0.2744
- Accuracy Unlabeled: nan

- Accuracy Stamp: 0.2744
- Iou Unlabeled: 0.0
- Iou Stamp: 0.2744

References:

SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers (https://arxiv.org/pdf/2105.15203)

How Transformers Work (https://towardsdatascience.com/transformers-141e32e69591)

Segformer Transformer (https://huggingface.co/docs/transformers/en/model_doc/segformer)

StaVer Dataset (https://www.kaggle.com/datasets/rtatman/stamp-verification-staver-dataset)