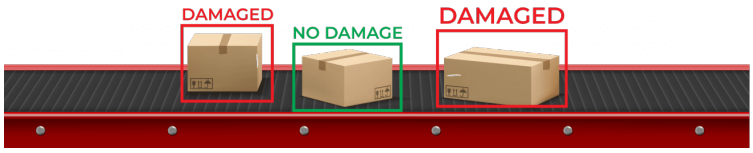
Parcel Damage Classification

William Swain, Helen Zhu, Zain Mirza

Virginia Polytechnic Institute and State University

Abstract



We set out to solve the problem of identifying damaged packages in warehouses. We explored various CNNs and found that the VGG model with a custom top layer resulted in the best performance, with a decently high accuracy, and a very high recall.

Introduction

- The rise of e-commerce places an increased demand on the parcel delivery industry. With the increased flow of parcels, distinguishing between good parcels ready for delivery and damaged parcels becomes difficult.
- Damaged parcels, which can be torn, leaking, crushed, incinerated, or folded, should not be sent out for delivery. Quickly and correctly classifying these parcels as ready for delivery or not is crucial for a company's efficiency.
- In this project, we applied several different models: ResNet50, VGG19, visualBERT, and Yolov5. We then compared them on several metrics to detemrine the best model for the problem of parcel classification..

Dataset

- The dataset comprises of a total of 1000 images, 500 images being images of packages ready to ship, 100 being crushed packages, 100 being folded packages, 100 being burned packages, 100 being leaking packages, and 100 being ripped packages.
- There is then 500 images of packages ready for ship and 500 images of packages that should not be shipped, allowing for a balanced dataset for binary classification.
- Splitting up the packages not ready to ship into its five subcategories allows for a more robust dataset that includes the many ways a package may be damaged.
- The images were collected using Google's iCrawler, which collects images directly from Google Images.
- Images are then cleaned by removing duplicates and irrelevant images, both of which are common when scraping from Google Images.

Methodology

The project follows the following pipeline:

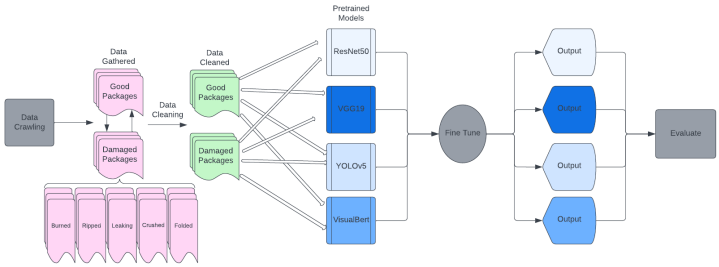


Figure 1: Project Pipeline

- For ResNet50 and VGG19, the top layers of the model is removed the rest of the base model layers have their weights frozen. Two dense layers and a global pooling layer are added on top of the model. These models are then trained using the cleaned custom parcel dataset.
- VisualBERT utilizes visual bounded region embeddings on top of the standard BERT model components. Visual, segment, and position components are summed, then passed to a multi-layer transformer. Self-attention enables the discovery of implicit relationships between language and visual features.
- YOLO v5 is a single-stage object detector with three essential parts: the Model Backbone extracts features, and in YOLO v5, the CSP - Cross Stage Partial Networks are used for this purpose. The Model Neck generates feature pyramids for object scaling, thereby improving its ability to generalize. The Model Head performs final detection using anchor boxes and output vectors with class probabilities, objectiveness scores, and bounding boxes.

Results

Midterm Results				
Metric	ResNet50	VGG19	Yolov5	visualBERT
Accuracy	0.888	0.95	0.9032	0.465
Precision	0.875	0.9	0.80036	0.036
Recall	0.897	1.0	0.73545	1.0
Specificity	0.878	0.909	0.8667	0.454
F1-Score	0.886	0.947	0.76637	0.070

Table 1: Model Results

Figure 2: Output

Conclusion

- ResNet50 performed decently in every metric, but was outperformed by Yolov5 in terms of accuracy, visualBert in terms of recall, and VGG19 in every metric.
- While YOLOv5's recall may not be high in general, it performs well in detecting good and bad boxes when fed with images of real-life boxes.
- The visualBERT model showed very poor performance and further investigation is need to determine the cause. We suspect dataset size and model generalizability are possible causes.
- From the table, VGG19 outperforms the other models in every evaluation metric. It is clear that VGG is the best model for classifying parcels.

References

- B. H. Team, "The Hidden Costs of Packaging Damage," BlackHawk Industrial Supply - The \#1 Choice For Manufacturers, 02-Jun-2020. [Online]. Available: https://www.bhid.com/hidden-costs-packaging-damage. [Accessed: 21-Mar-2023].
- Bird Species Classification from an Image Using VGG-16 Network: Proceedings of the 7th International Conference on Computer and Communications Management. ACM other conferences. [Accessed: 21-Mar-2023].
- Kim et al. (2022, March 3). Image classification of parcel boxes under the underground logistics system using CNN mobilenet. MDPI. [Accessed: 21-Mar-2023].
- K. Simonyan et al. (2015, April 10). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv.org. [Accessed: 21-Mar-2023].
- L. Li et al. (2019, August 9). VisualBERT: A Simple and Performant Baseline for Vision and Language. arXiv.org. [Accessed: 21-Mar-2023].