Vehicle Classification Using Computer Vision for KAR Global

UChicago MScA Capstone Project

Yue Wu, Chuyu Chen, Hsuan-Hao Lin, Jack Gao

March 11, 2022

Agenda

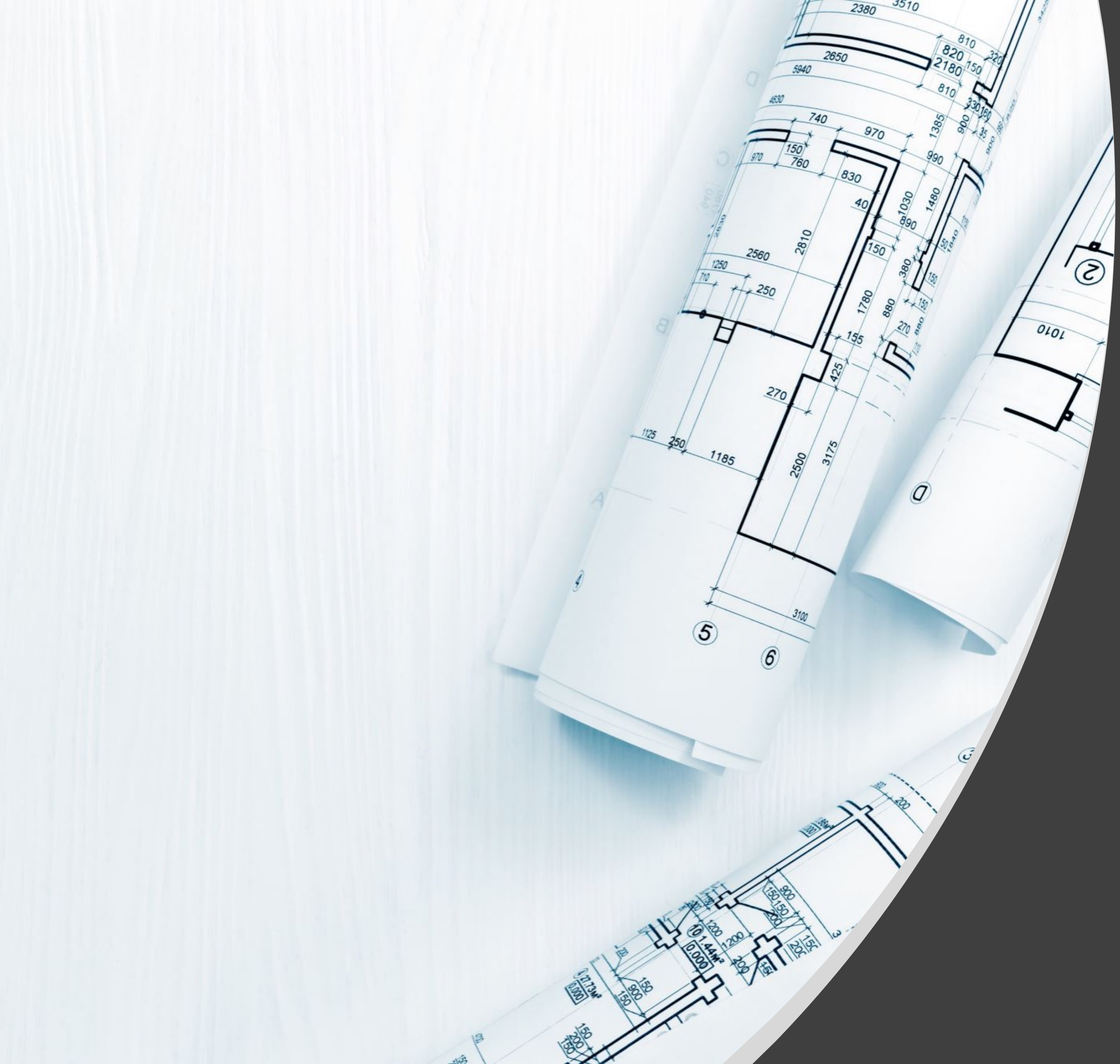
1 Project Overview

2 Data & Feature Engineering

3 Methodology & Findings

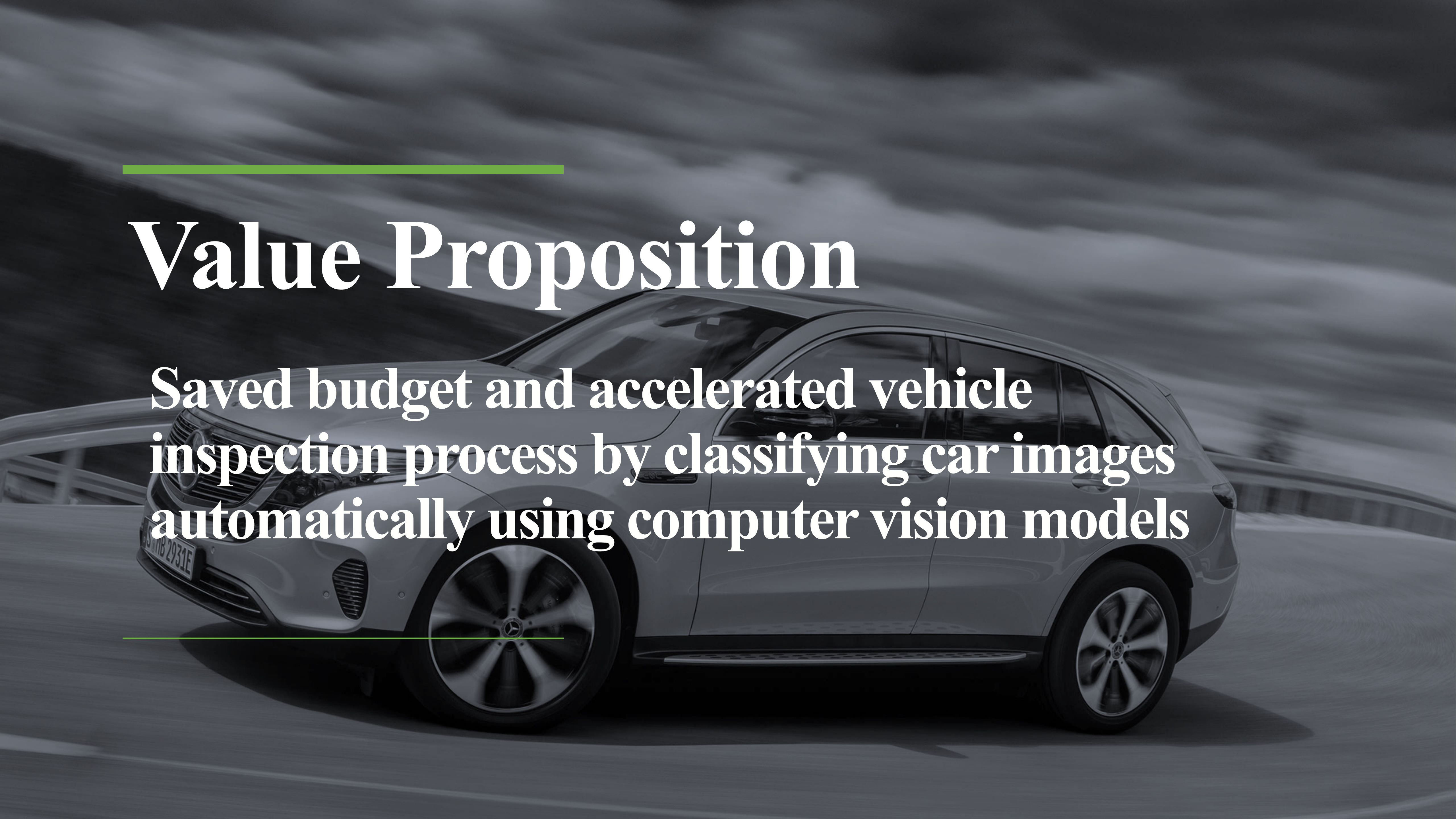
4 Conclusion & Demo

KAR GLOBAL



1. Project Overview

- Value Proposition
- Client Background
- Problem/Opportunity Statement
- Goals & Challenges



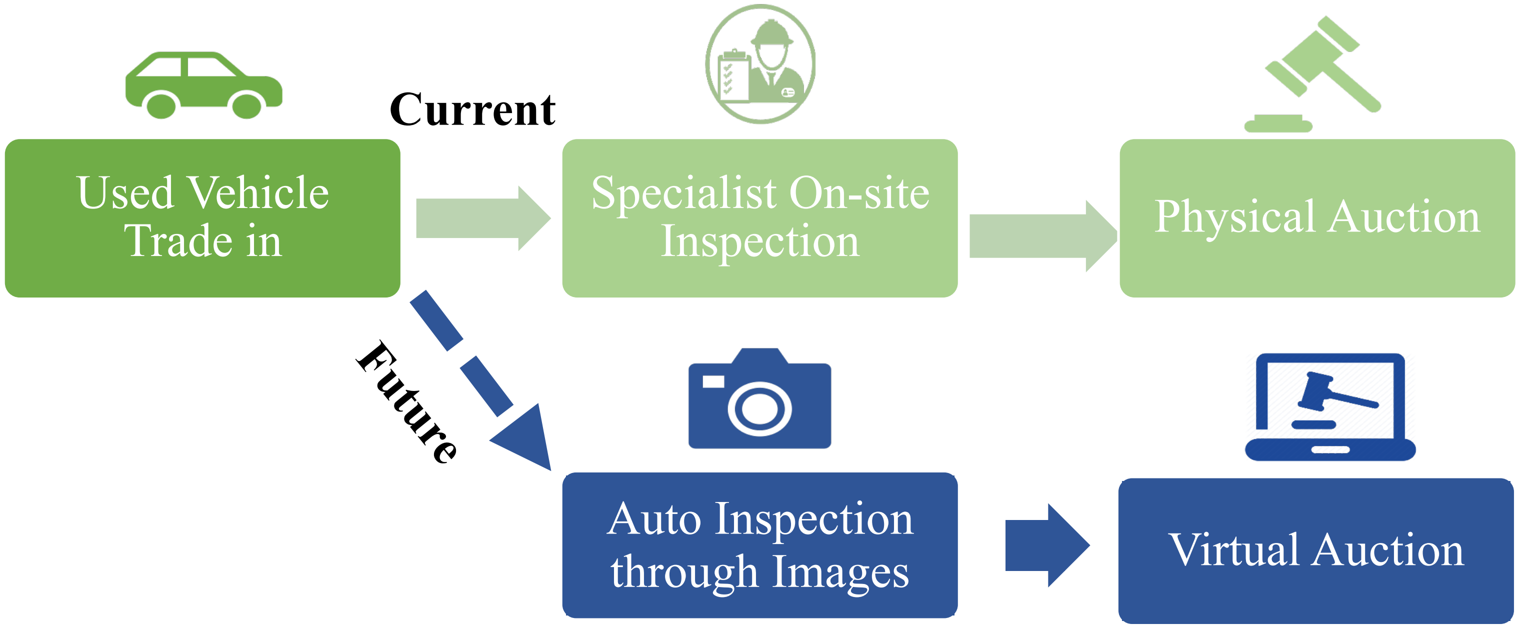
Value Proposition

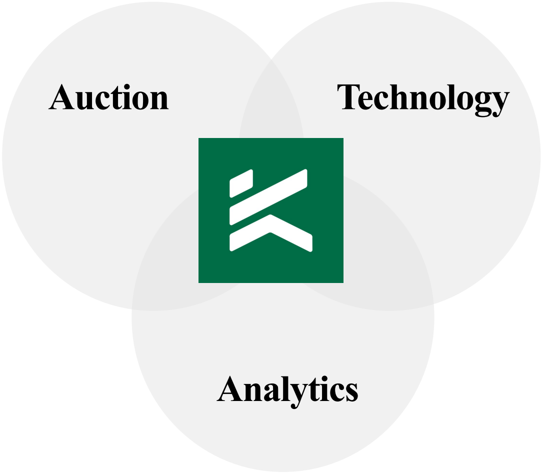
Saved budget and accelerated vehicle inspection process by classifying car images automatically using computer vision models

Client Background

Overview of KAR Global

- KAR Global is a leading provider of used vehicle auctions and a company that provides remarketing solutions to sellers and buyers
- KAR Global is a company that integrates technology, analytics, and auction to leverage data and accelerate remarketing





Source: KAR Global annual report

Problem/Opportunity Statement

Business Problems

- Mislabeling and inefficiency from manual classification process
- Increasing difficulty of conducting car inspection at physical auction site due to COVID
- Increasing demand of real-time vehicle report data to help identify the cars that the customers want

Opportunities

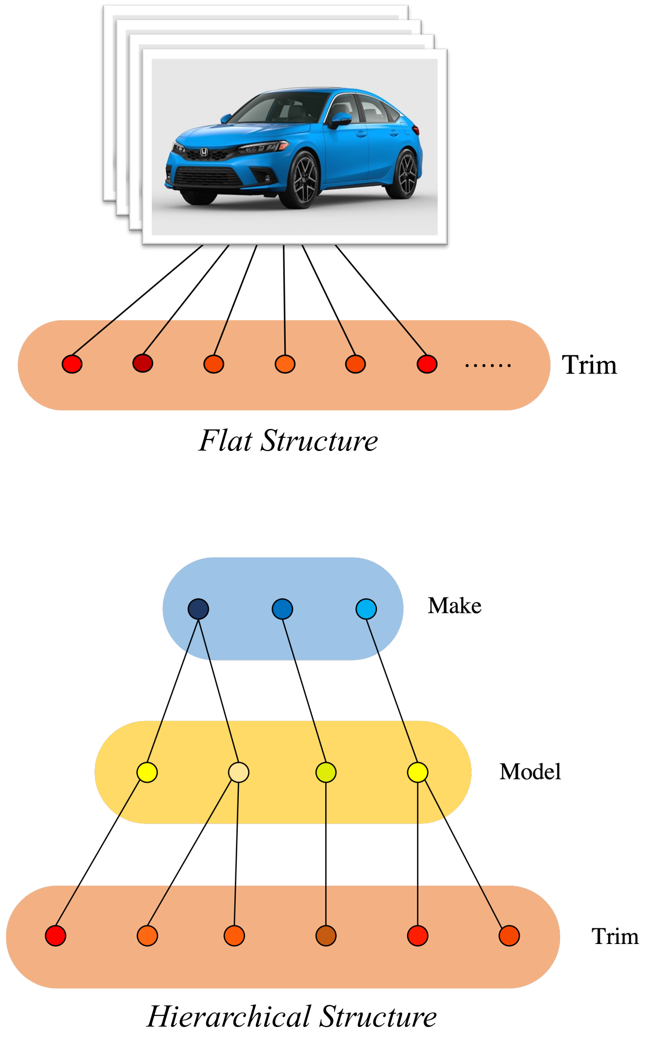
- Use computer vision models to classify car images
- Automate entire car inspection process and potentially move it upstream
- Accelerate the classification process and save budget

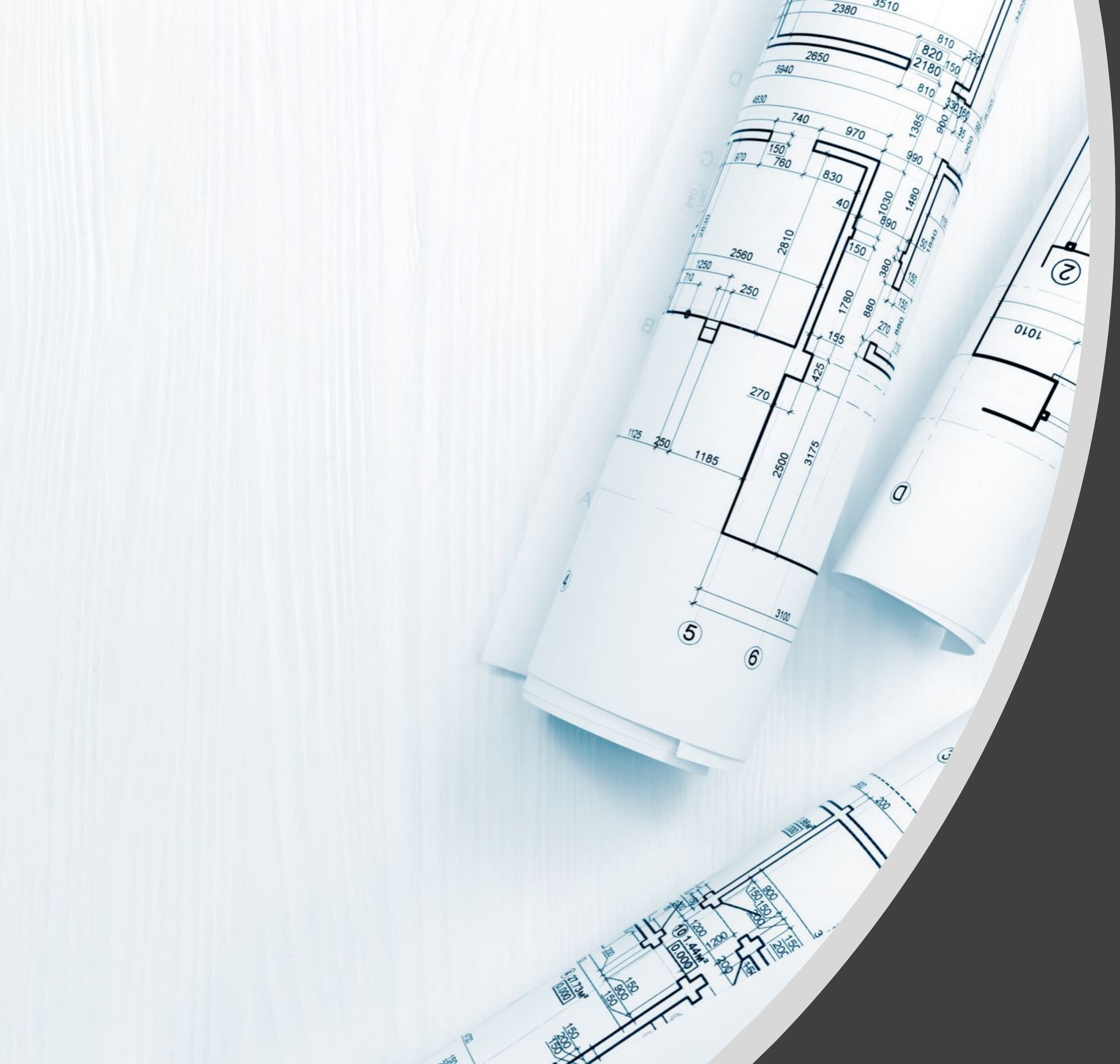
Goals & Challenges

Phase 1: Research, Exploratory Analysis & Data Cleaning
- **Goal:** Figure out potential modeling architectures, and get dataset ready for training
- **Challenge:** Massive data cleaning for trim level; poor image quality

Phase 2: Explore Flat and Hierarchical Structure
- **Goal:** Use hierarchical CNN to classify make, model and trim of a vehicle
- **Challenge:** Computing power constraint; multiclass classification with large number of categories; subtle image differences at trim level

Phase 3: Integrate Model Into Production Pipeline
- **Goal:** Facilitate seamless model integration





KAR GLOBAL

2. Data & Feature Engineering

Vehicle Image Dataset

68
Vehicle
YEAR

6,034
Vehicle
MODEL

10 M+
Vehicle Images

936
Vehicle
MAKE

22,830
Vehicle
TRIM

Image Angles
1. Front 2. Rear 3. Left front 4. Left rear 5. Right front 6. Right rear

Data Cleaning

1 Initial Cleaning

- Drop NULL
- Unify image angle notation
- Limit the **YEAR** range to 2011-2021

2 Label & URL Cleaning

- Text Normalization
- Punctuation / Whitespace Removal
- Label Stemming

Feature Engineering

1. Prioritization

Shortlist make, model, and trim to **top labels**

2. Consolidation

Consolidate model and trim labels that are **similar or almost identical**

3. Transformation

Transform remaining make and model that are outside of top labels to "**OTHER**"

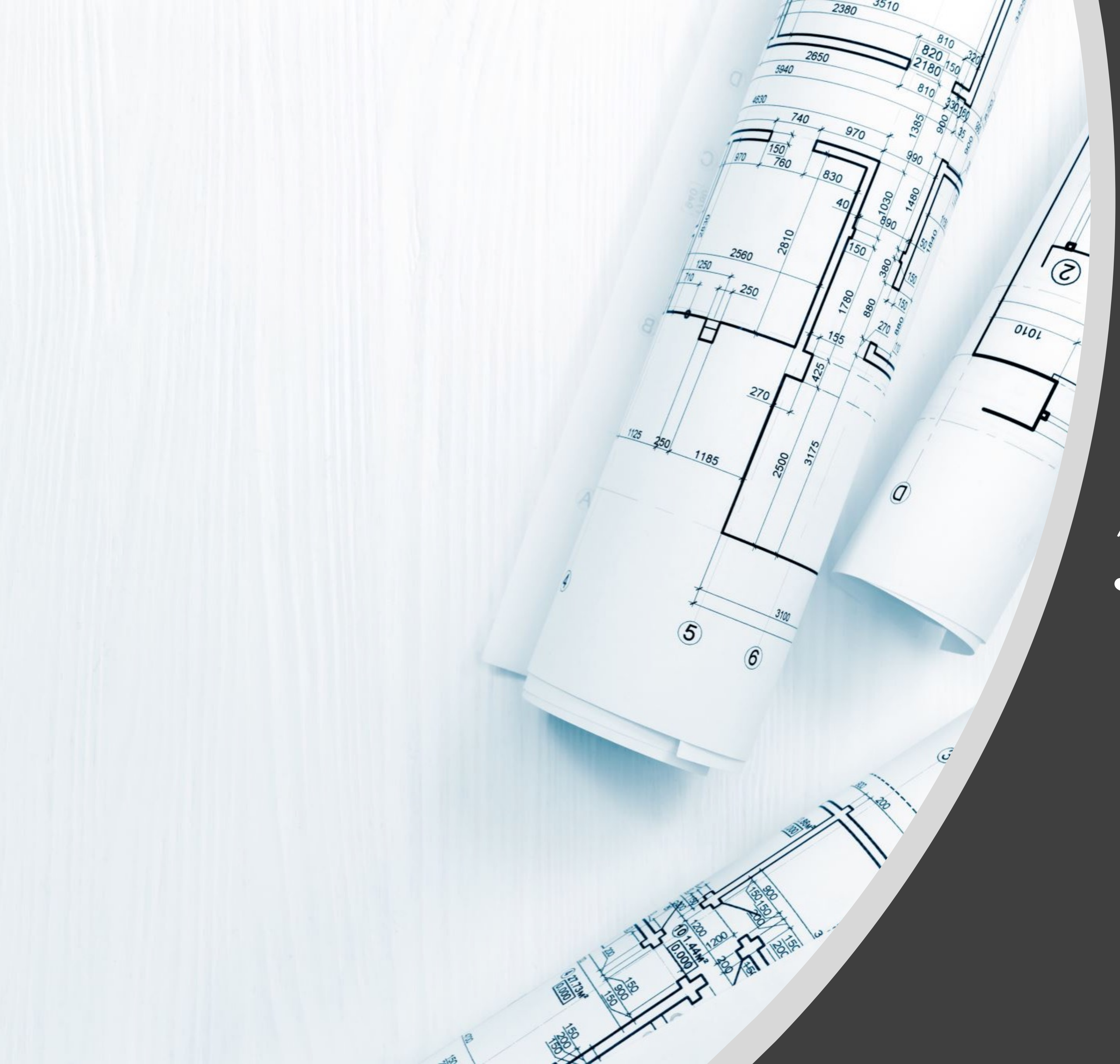
Images from **Top 20** MAKEs make up 90% of original dataset

Similar Labels	Consolidated
• CIVIC SDN • CIVIC Sedan	• CIVIC Sedan

21 (20 + 1)
MAKEs

119
MODELs

361
TRIMs



KAR GLOBAL

3. Methodology & Findings

Methodology (I)

Experiment I: Compare Model Performance Trained on Different Image Angle

Input Image Angle

Front	Rear	Front + Rear

Findings (I)

Experiment I: Compare Model Performance Trained on Different Image Angle

Test Accuracy for different CNN configuration and image angles

Image Angle	MobileNet v2	VGG-16	DenseNet 121	ResNet 152	EfficientNet B0	Average
★ Front	**0.951**	**0.794**	**0.957**	**0.921**	**0.972**	**0.919**
Rear	0.734	0.658	0.870	0.855	0.937	0.811
Front + Rear	0.816	0.682	0.754	0.670	0.931	0.771

Takeaway:

- **Front** images generate the best result compared to Rear or Front-Rear combination across all CNN architectures consistently

Methodology (II)

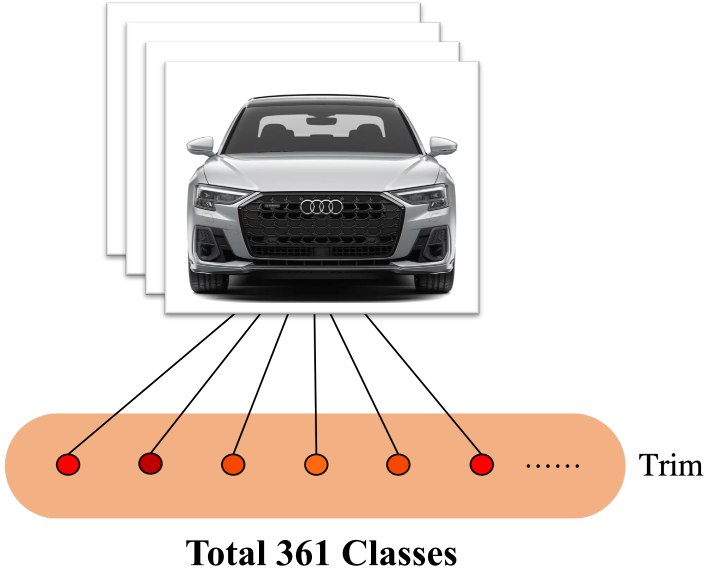
Experiment II: Flat 1-Layer Model for Trim Classification

CNN Architecture

MobileNet v2
VGG-16
DenseNet 121
ResNet 152
EfficientNet B0

Input Image Angle

Front



Finding (II)

Experiment II: Flat 1-Layer Model for Trim Classification

Test Accuracy for different CNN configuration

	MobileNet v2	VGG-16	DenseNet 121 ★	ResNet 152	EfficientNet B0
Test Accuracy	0.733	0.681	**0.756**	0.733	0.691

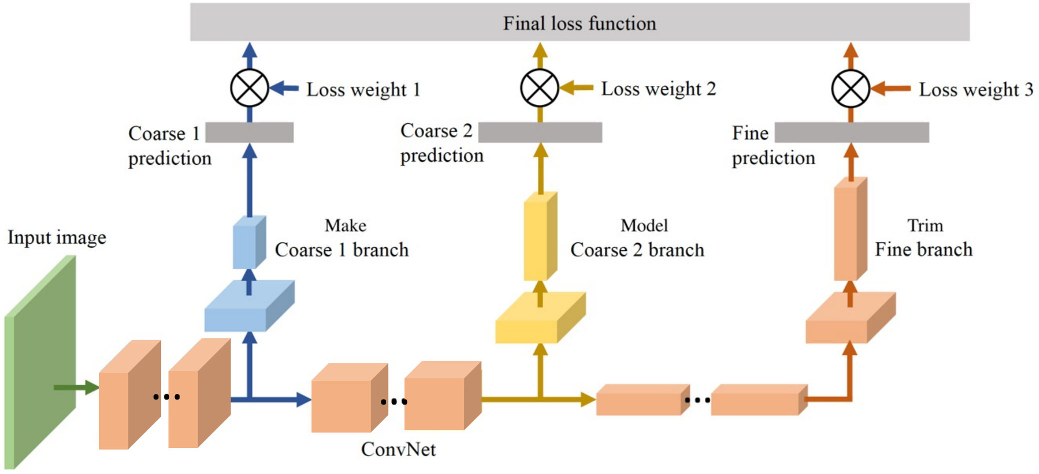
Takeaway:

- **DenseNet 121** outperformed other CNN architectures in classifying vehicle Trim for flat 1-layer model

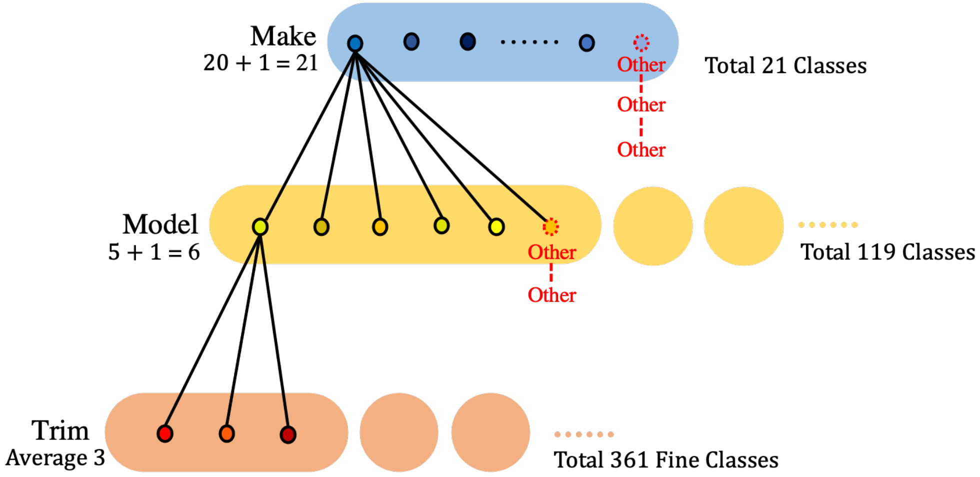
Methodology (III)

Experiment III: Hierarchical 3-Layer Model for Make, Model, and Trim Classification

- **B-CNN** consists of a main convolution workflow just like a normal CNN, with several branch networks attached to it to predict the corresponding level of categories
- **DenseNet 121** as main network in the B-CNN suggested by prior experiment
- **Branch Training Strategy** with low-level weights are activated earlier than high-level weights



B-CNN Architecture



Data Structure of Hierarchical Classification

Finding (III)

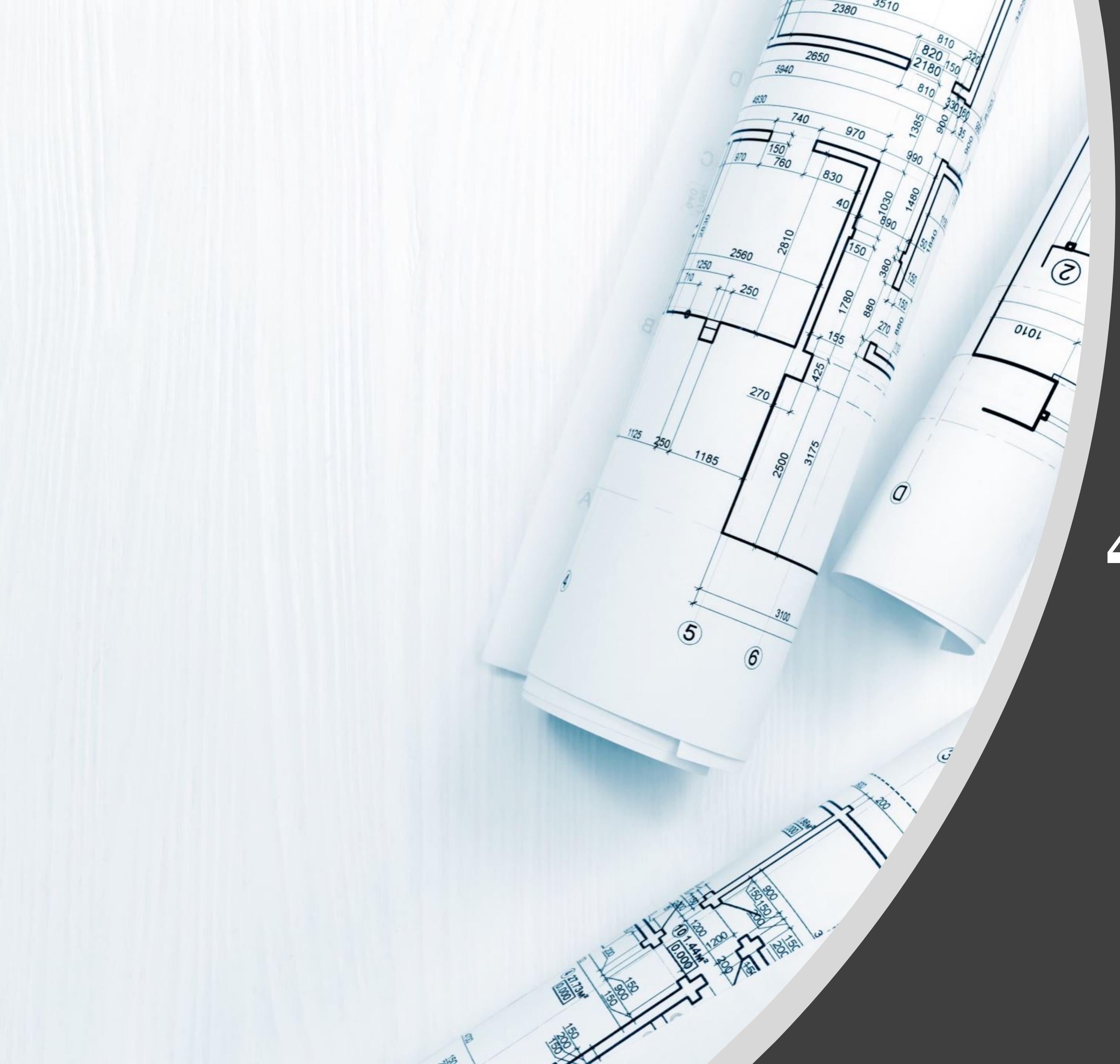
Experiment III: Hierarchical 3-Layer Model for Make, Model, and Trim Classification

Test Accuracy Comparison between DenseNet 121 B-CNN and Flat

Model	DenseNet 121 B-CNN	DenseNet 121 Flat
Layer 1 (Make)	**0.980**	-
Layer 2 (Make - Model)	**0.968**	-
Layer 3 (Make – Model - Trim)	**0.738**	0.756

Takeaway:

- **DenseNet 121 B-CNN** achieved outstanding accuracy for Make and Model
- Regarding **Layer 3 accuracy**, flat structure is slightly better than hierarchical structure



KAR GLOBAL

4. Conclusion & Demo

Conclusion

1. Front Images generate the best accuracy

Compared to Rear, and Front & Rear combination, **Front images** tend to produce the best prediction accuracy across all architectures

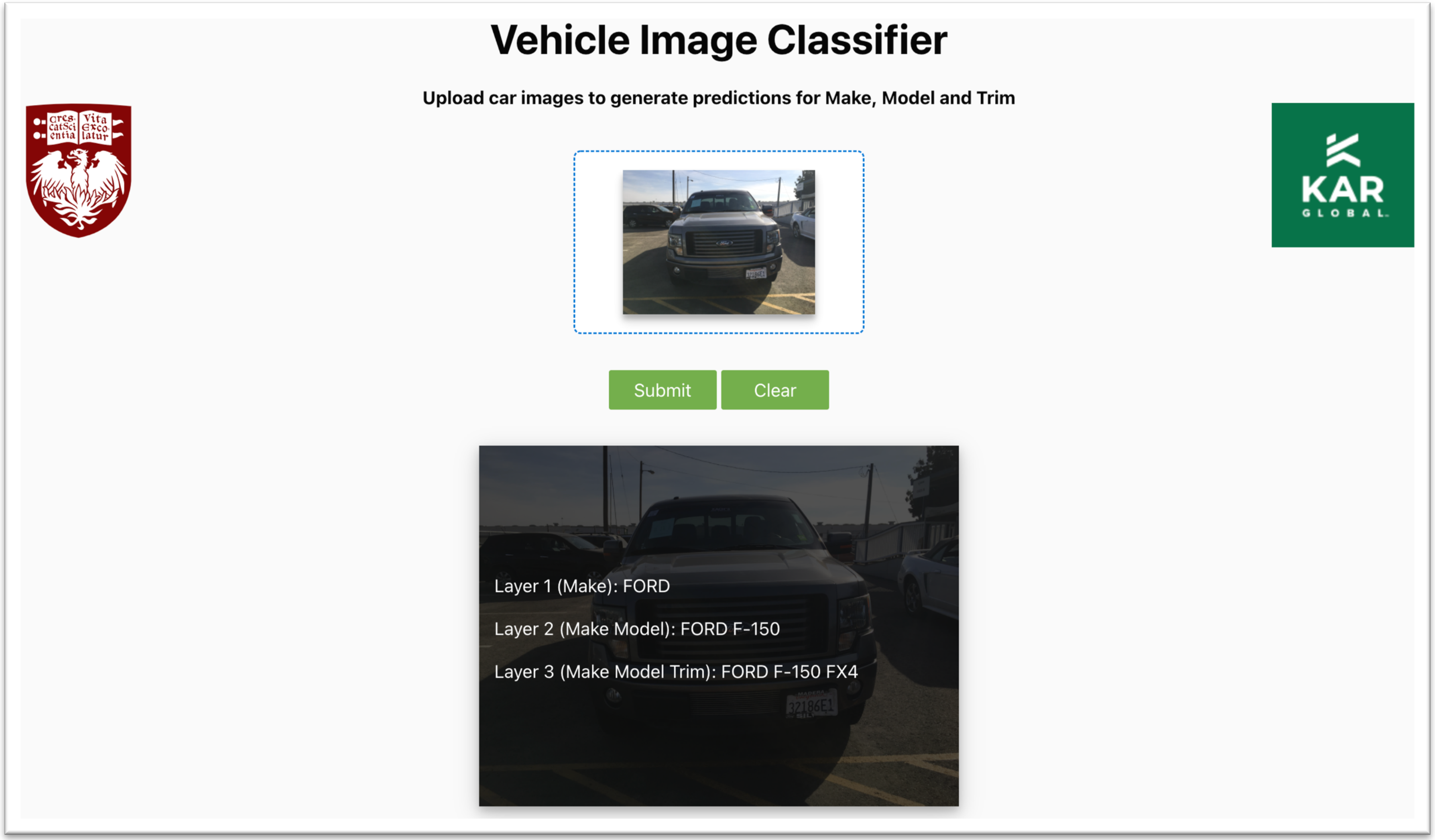
2. DenseNet 121 performs the best for flat structure

Among existent CNN models, **DenseNet 121** claims the title for doing one-layer classification by achieving a test accuracy of **0.756**.

3. Hierarchical structure offers additional benefits

Although the hierarchical structure has slightly lower trim accuracy, it is still better than flat structure due to its **high efficiency** and **interpretability**.

Demo

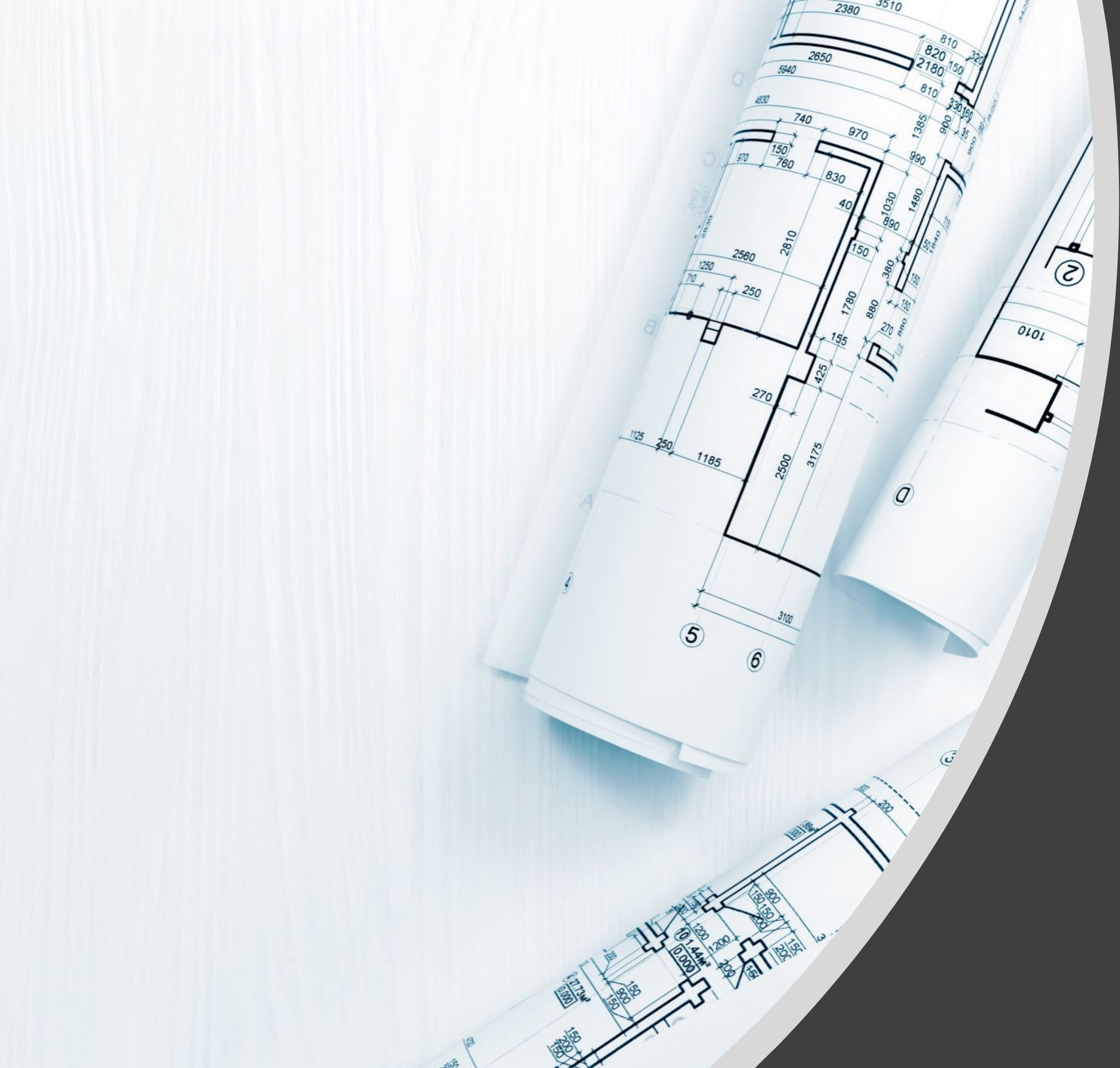

Vehicle Image Classifier
Upload car images to generate predictions for Make, Model and Trim
Submit
Clear
Layer 1 (Make): FORD
Layer 2 (Make Model): FORD F-150
Layer 3 (Make Model Trim): FORD F-150 FX4

Special Thanks
To

Bryan Binotti

Ahad Beykaei

Dr. Arnab Bose



Appendix

Training Process of DenseNet 121 B-CNN

Loss and Accuracy of First Layer (c1), Second Layer (c2), and Third Layer (c3) of the B-CNN Classification Model