

CarMatch

A singular car buying solution

Thomas Noriega

Car Buying is for Experts



CARVANA

Find My Car

Trade-In

Financing

Learn More ▾

Sign In | [Create Account](#)

THE **NEW** WAY TO BUY A CAR

Delivery to Your Door | 7-Day Test Own

HOW IT WORKS



SEARCH OUR 9,950 CARS

Select **BODYSTYLE(S)**



SEDAN



SUV



PICKUP



MINI-VAN



HATCHBACK



CONVERTIBLE



COUPE



WAGON

Select **PRICE RANGE**

< \$15K

\$15K-20K

\$20K-25K

\$25K-30K

\$30K-35K

> \$35K

FIND MY NEXT CAR


Car Buying is for Experts

Need to know **exactly** what you want

THE **NEW** WAY TO BUY A CAR

Delivery to Your Door | 7-Day Test Drive

SEARCH OUR 9,950 CARS

 I want... (ex. Audi A5)

Select BODYSTYLE(S)



SEDAN



SUV



PICKUP



MINI-VAN



HATCHBACK



CONVERTIBLE



COUPE



WAGON

Select PRICE RANGE

< \$15K

\$15K-20K

\$20K-25K

\$25K-30K

\$30K-35K

> \$35K

FIND MY NEXT CAR

Car Buying is for Experts

Need to know **several constraints**

Select **BODYSTYLE(S)**



SEDAN



SUV



PICKUP



MINI-VAN



HATCHBACK



CONVERTIBLE



COUPE



WAGON

Select **PRICE RANGE**

< \$15K

\$15K-20K

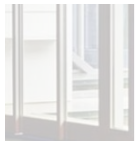
\$20K-25K

\$25K-30K

\$30K-35K

> \$35K

FIND MY NEXT CAR



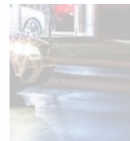
HATCHBACK

CONVERTIBLE

COUPE

WAGON

FIND MY NEXT CAR



CarMatch

Just needs a *single, understandable* input

users don't need to be experts to browse
listings

websites can reduce friction and increase
customer conversion

Inception 2D Neural Network

Modified and progressively fine-tuned to classify
196 different car classes

- No feature extraction
- Minimal pre-processing
- Could be trained on a small dataset

Dataset

Each class a unique combination of:



MAKE: Dodge
MODEL: RAM 3500
YEAR: 2009

Approximately **80 photos/class** (~16,000 photos total)

http://ai.stanford.edu/~jkrause/cars/car_dataset.html

Multiple Angles



Variable Quality



Validation on Test Set

Top-1 accuracy: 77.1%

Top-5 accuracy: 94.8%

Validation on Test Set

Top-1 accuracy: 77.1%

Top-5 accuracy: 94.8%

So if the user just wants to identify the uploaded car they almost certainly will

Validation on Test Set

Top-1 accuracy: 77.1%

Top-5 accuracy: 94.8%

So if the user just wants to identify the uploaded car they almost certainly will

But results are most useful if they also ***look like the submitted photo***

Input: sedan

output: sedans



Input: sedan



output: sedan-like



Similarity Validation

Incorrect ID: 5.2%

Similarity Validation

Incorrect ID: 5.2%

Correct ID



Similarity Validation

Incorrect ID: 5.2%

Correct ID AND styles



Similarity Validation

Incorrect ID: 5.2%

Correct ID AND styles: 15.8%



Similarity Validation

Incorrect ID: 5.2%

Correct ID AND styles: 15.8%

Correct ID AND wrong styles: 79.0%



How to improve the results

Instead of absolute top-5 ranking

Algorithm that balances

- 1) most represented **body-styles**
and
- 2) the **highest confidence** results

Improved results by 4x

Correct ID AND styles:

New algorithm: **71.4%** (vs. 15.5%)

Correct ID AND wrong styles:

New algorithm: **21.5%** (vs. 79.0%)

4x
improvement

Incorrect ID:

New algorithm: **7.1%** (vs. 5.2%)

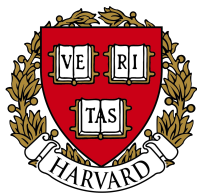
1.4x
compromise

And that's how CarMatch

gives customers more results, more easily

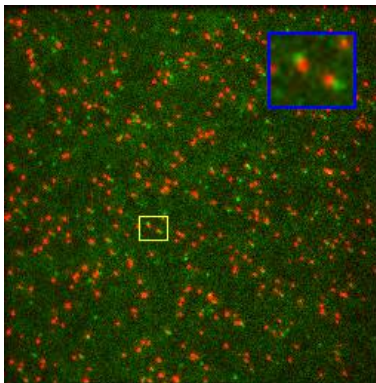
helps websites sell more cars

Thomas Noriega



UCSF

University of California
San Francisco



SUPPLEMENTAL SLIDES

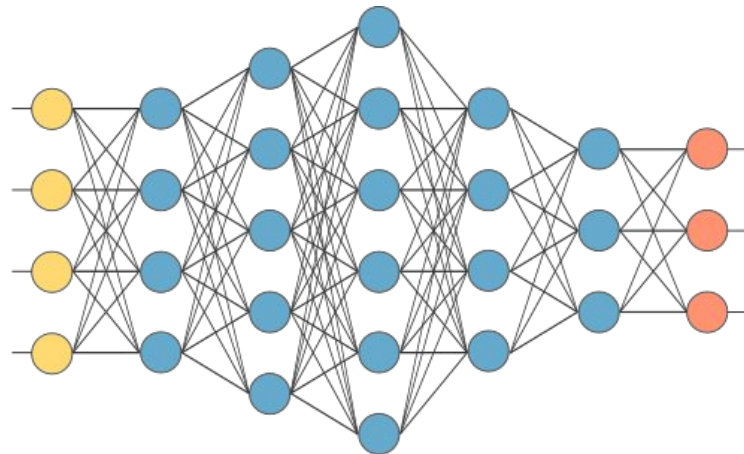
Training



training set

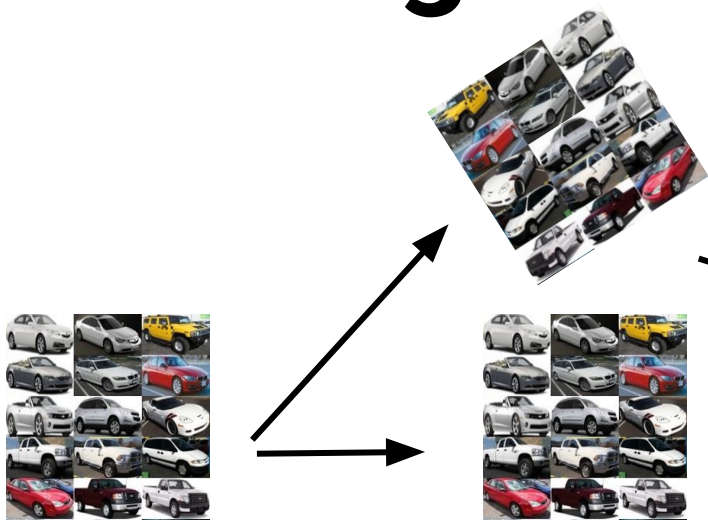


in-memory
augmentation



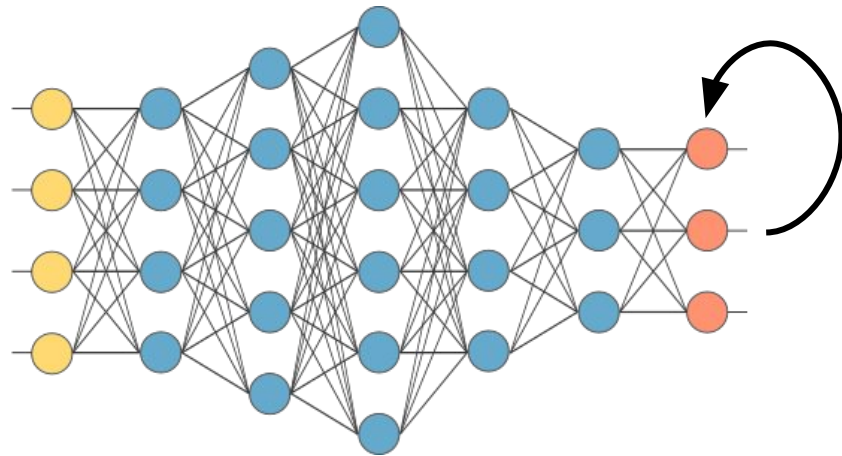
Pre-trained GoogleNet (Inception)
2D convolutional neural network

Training



training set

in-memory
augmentation



2D convolutional neural network
add/train new prediction layer

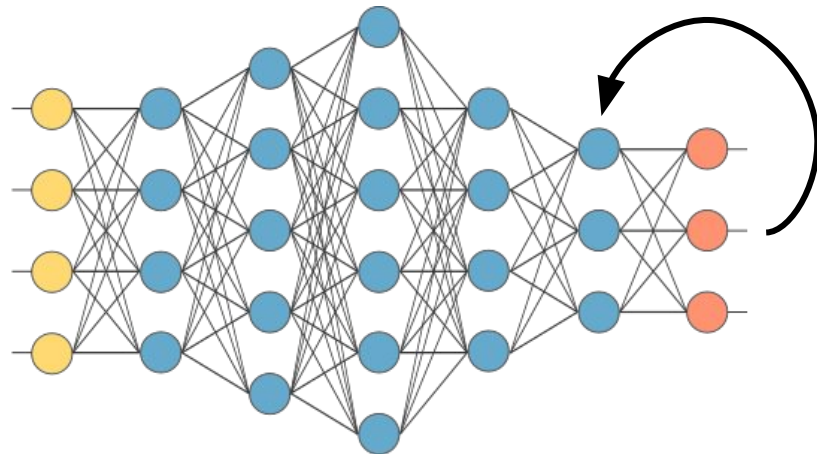
Training



training set



in-memory
augmentation



2D convolutional neural network
progressive fine-tuning

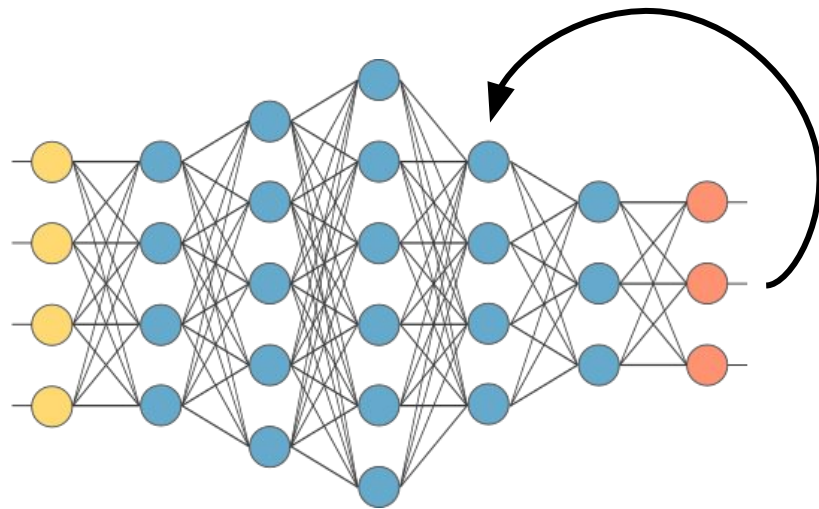
Training



training set



in-memory
augmentation



2D convolutional neural network
progressive fine-tuning

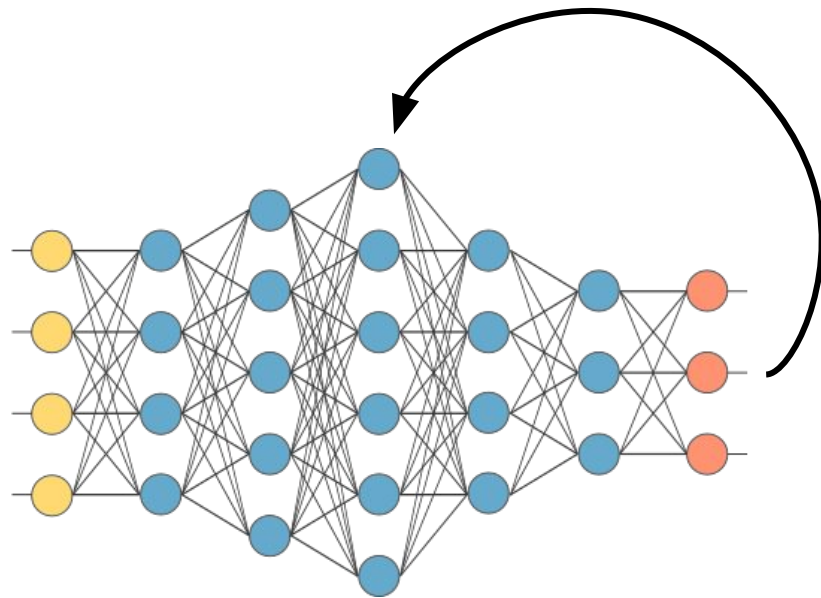
Training



training set

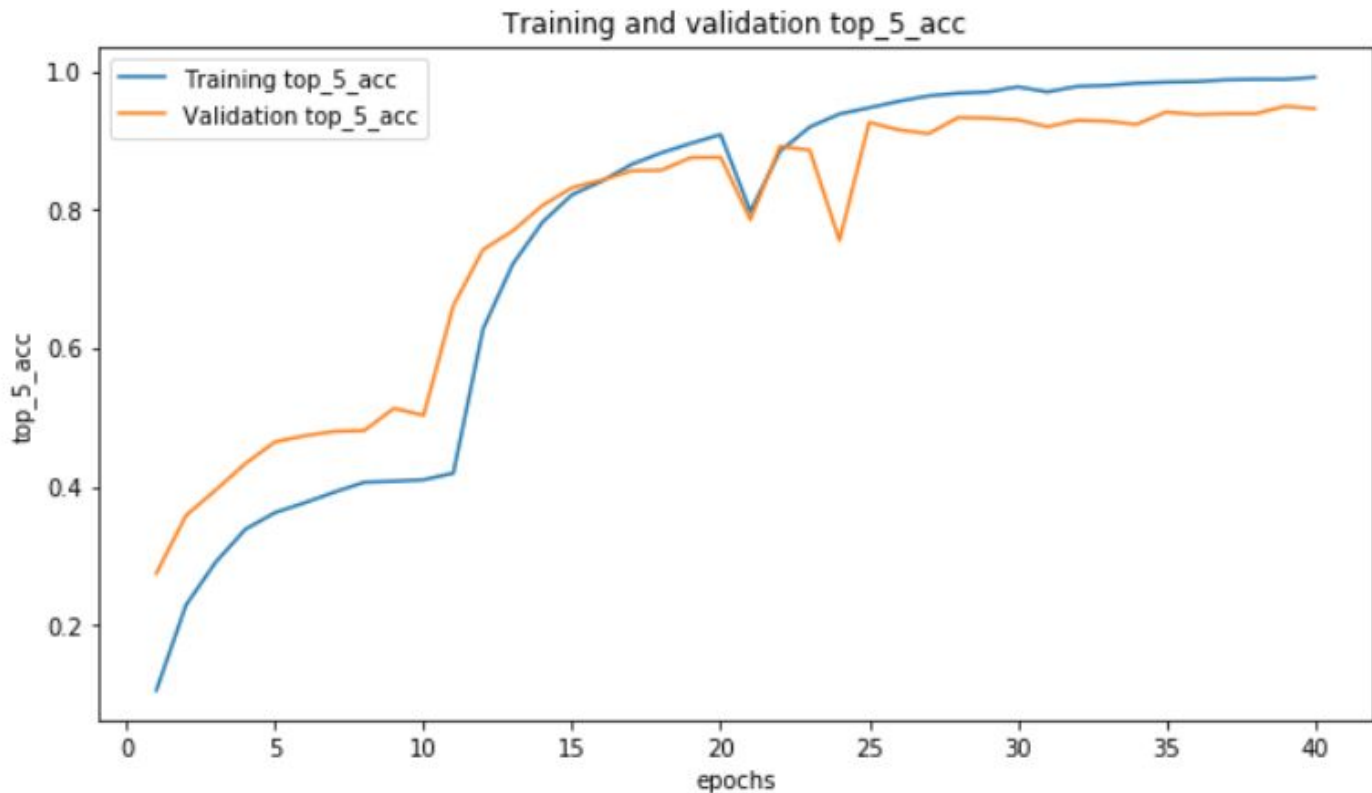


in-memory
augmentation

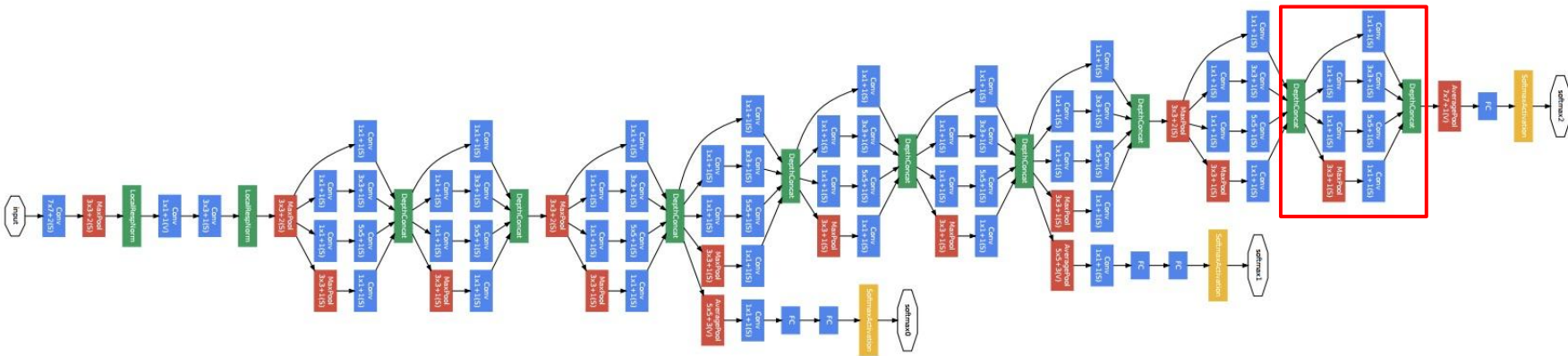
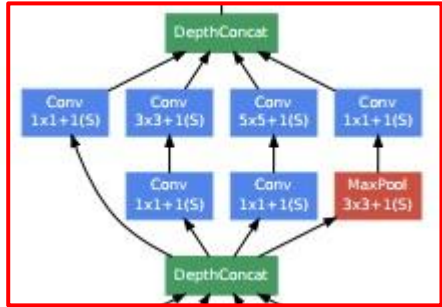


2D convolutional neural network
progressive fine-tuning

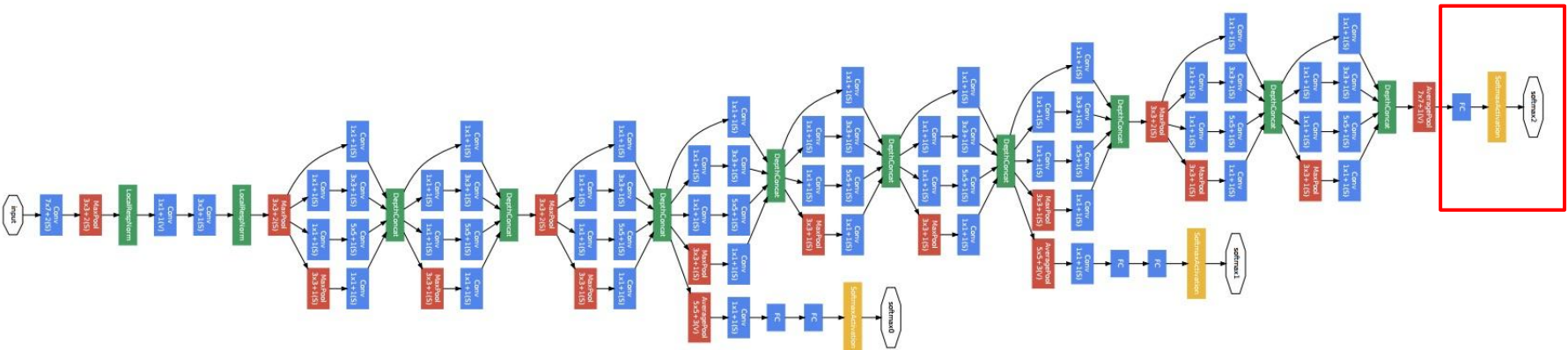
Training/Validation Curves



Inception Model



Modify



Why such high wrong style results?

Top-5 ignores probability assignments:



0.85



0.10



0.02



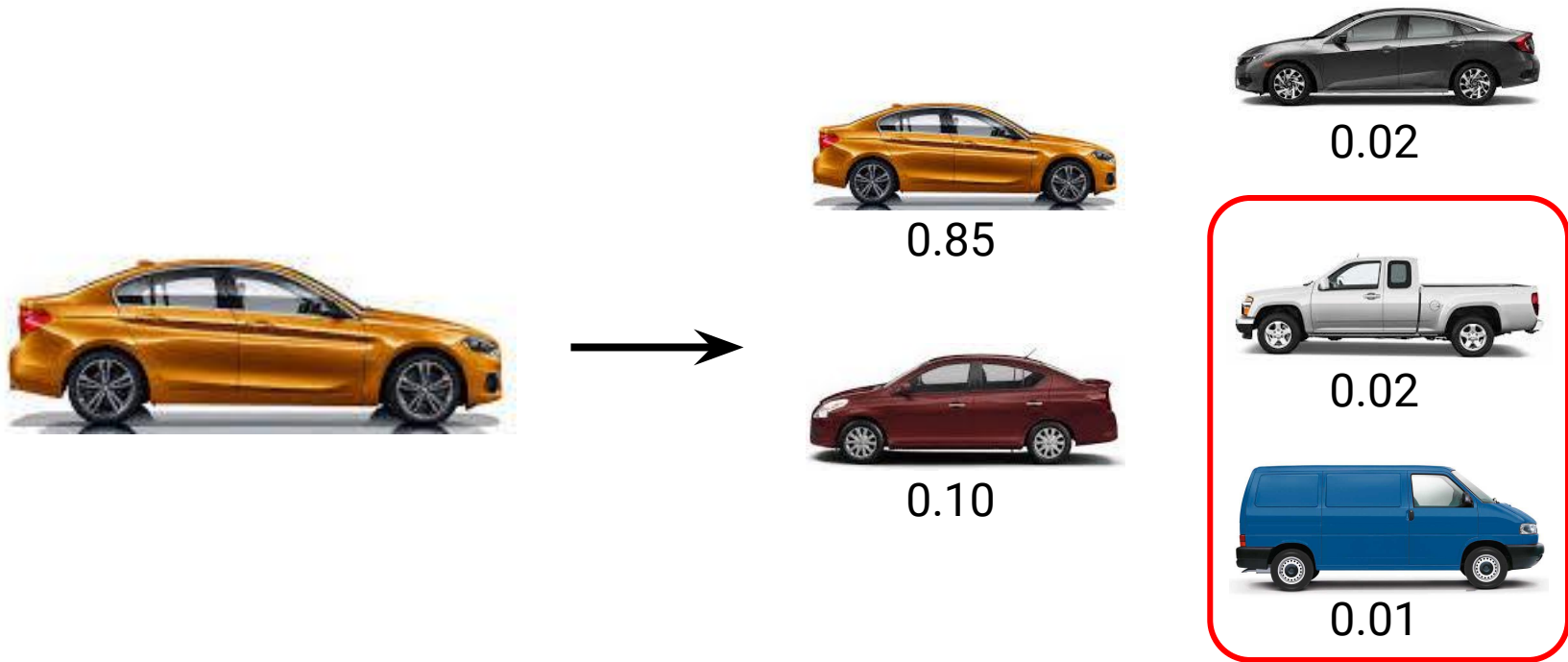
0.02



0.01

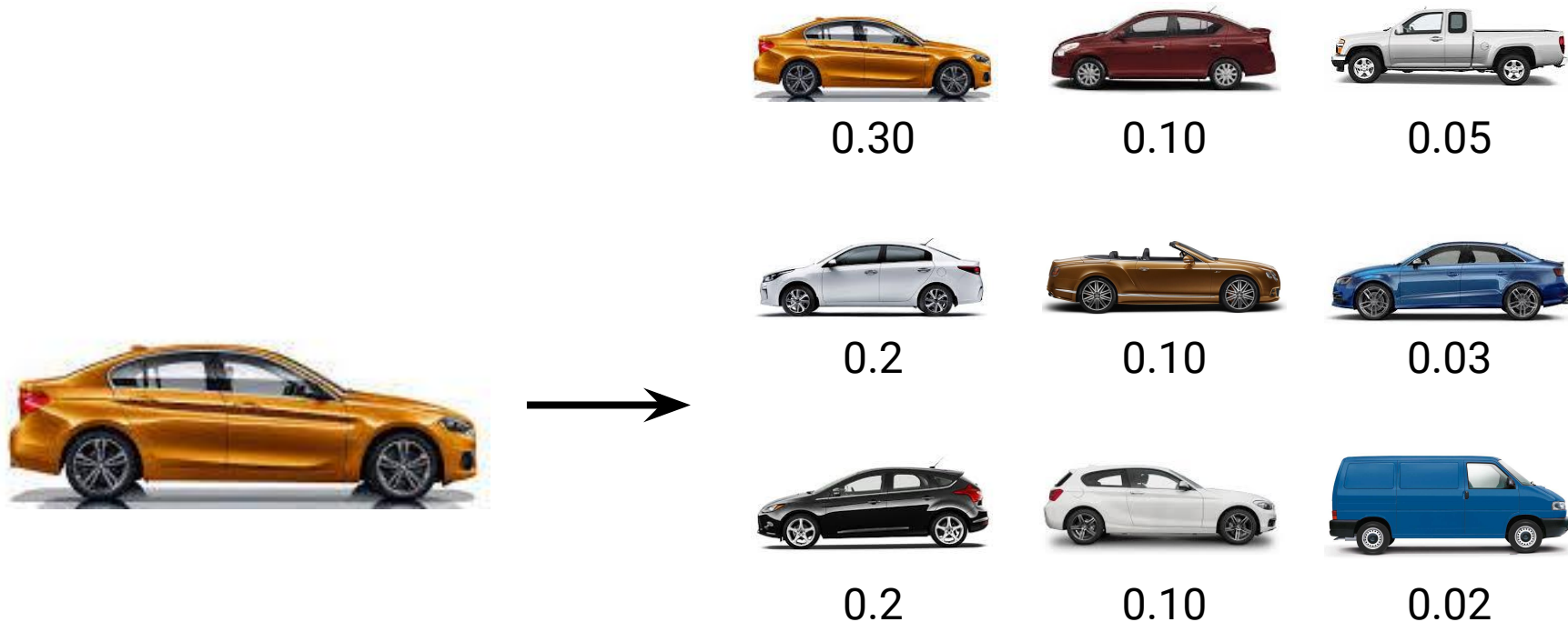
Why such high wrong style results?

Top-5 ignores probability assignments:



How to improve the results

1) Start with all predictions



How to improve the results

- 1) Start with all predictions
- 2) Consider predictions with 95% confidence



0.30



0.10



0.05



0.2



0.10



0.03



0.2



0.10



0.02

How to improve the results

- 1) Start with all predictions
- 2) Consider predictions with 95% confidence
- 3) Limit to top-2 body-styles



0.30



0.10



0.05



0.2



0.10



0.03



0.2



0.10



0.02

How to improve the results

- 1) Start with all predictions
- 2) Consider predictions with 95% confidence
- 3) Limit to top-2 body-styles
- 4) Present top-5 remaining results



0.30



0.10



0.05



0.2



0.10



0.03



0.2



0.10



0.02