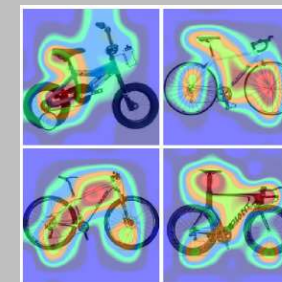


The Price is Right: Predicting Prices with Product Images

(2018) Steven Chen / Edward Chou / Richard R. Yang

<https://arxiv.org/pdf/1803.11227>



どんなもの？

自転車と自動車の画像からそれぞれの価格を推定するregressionタスクにCNNを用いた。ベースラインモデルである多クラスSVMと線形モデルに対して同等～高い精度が出た。VGGとMobileNetによる転移学習よりも制度の高いネットワーク構造PriceNetReg CNNを提案している。また、CNNモデルに関して、画像のどの部分が予測に効いているかをVisualizationすることにより示している。

どうやって有効だと検証した？

実際に自動車と自転車の画像と価格をクロールして、それらに対してベースラインモデルと転移学習によるCNN、著者らが提案しているPriceNetReg CNNそれぞれに対してRMSE, MAE, R², Precision Recall F1-scoreを比較して有効だと検証した

技術の手法や肝は？

簡潔な構造のPriceNetReg CNNによって、ベースラインモデルと転移学習モデルをアウトパフォーム

議論はある？

Saliency Map, Class Activation Mapという手法でどの部分からCNNが判断を行っているかということを見ている。画像の輪郭部分と自動車だとロゴ・タイヤの部分・自転車だとハンドル・シート・補助輪・スポークの部分などが着目されている

先行研究と比べて何がすごい？

今までは、画像をカテゴライズするようなネットワーク構造が多い。それに対して、特に価格予測にフォーカスしてトレーニングが行われており、簡潔である

次に読むべき論文は？

CAMの論文：

R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In ICCV, 2017.

2018/04/14