Improved prediction of environment-based smoking risk from images of daily life

Abhishek Jadhav*, **Matthew M Engelhard***, **Jason A Oliver***, **Cynthia Conklin**†, **and F Joseph McClernon***

*Department of Psychiatry and Behavioral Sciences, Duke University Medical Center

†Department of Psychiatry, University of Pittsburgh Medical Center

Duke CfAST
Center for Addiction Science and Technology

1. Introduction

- **Smokers report increased craving and tobacco use when viewing environments where they habitually smoke compared to environments where they do not,** suggesting that these environments may contribute to lapses and relapse following a quit attempt.
- Our previous research has demonstrated that smoking environments can be recognized with a deep learning approach, by identifying which settings in images of daily life are predictors of smoking risk. The results suggested that the technique could be used to support **just-in-time adaptive interventions (JITAIs)**, or to **identify specific environmental cues that may confer risk for smoking and potentially other target behaviors**.
- We extend previous work by utilizing mobile-friendly architectures to **move towards real-time smoking risk prediction elicited by environments smokers encounter while maintaining performance** as well as training our model on an expanded cohort of over 300 smokers from the Durham, NC and Pittsburgh, PA areas.
- This framework for interpreting and predicting the influence of environments on target behaviors **provides a basis for environment-based interventions**, with broad applications in mental and physical health.

2. Participants and Methods

- As part of longstanding research on environments and smoking, 169 adult (18-55) smokers (>5/day) from the Durham, NC (N=106) and Pittsburgh, PA (N=63) areas photographed ≤4 of their **smoking environments** and ≤4 **nonsmoking environments**.
- These images (N=2903) were used to train a deep learning model that predicts the probability of each location type (smoking/nonsmoking), which may then be used to **approximate environment-based smoking risk**.
- Our classifier is comprised of a light CNN (MobileNetV2) and object detection framework (Single Shot Detector) for feature extraction, with an interpretable logistic regression model or multi-layered perceptron at the output. It was trained and evaluated via nested cross-validation with respect to patients (i.e. out-of-patient prediction).
- To contextualize model performance, results were compared with our previous research.

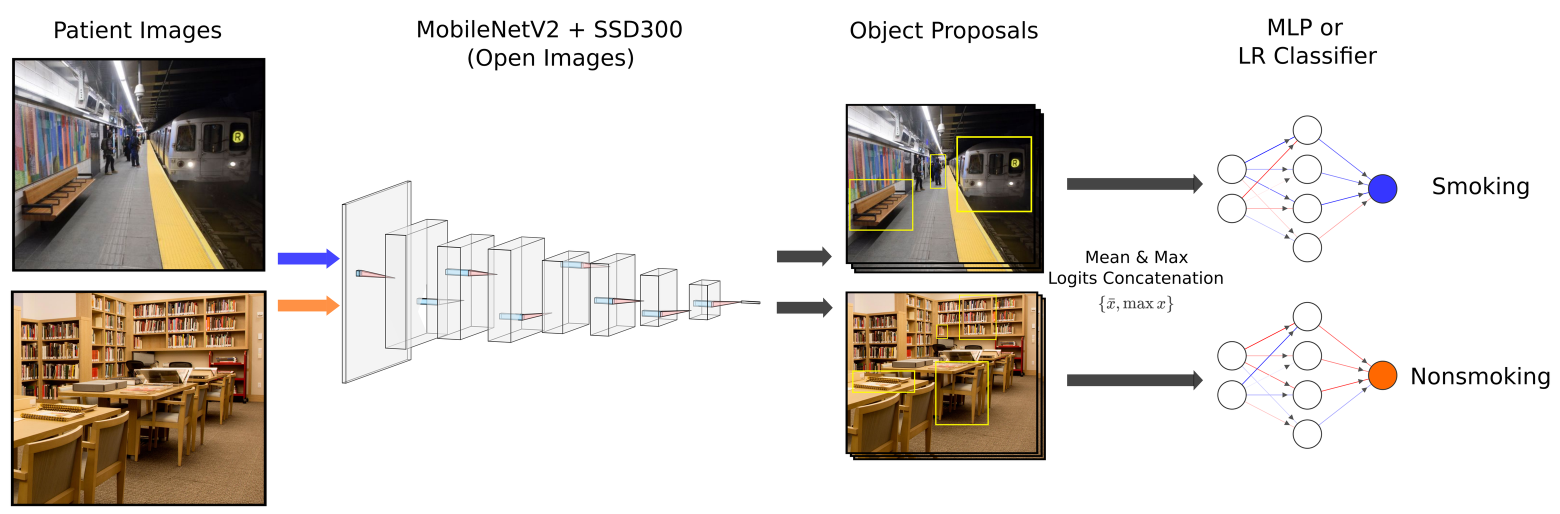

Figure 1: *Illustration of the classification model, which extracts image features using a MobileNetV2 base architecture paired with an SSD object detector, the maximum and average logits for each detection class were calculated across all object proposals. This was then fed to a multi-layered perceptron (or logistic regression) to classify the images as a smoking environment or nonsmoking environment.*

3. Results

- The logistic regression variant discriminated environment types with 0.816 AUC (74.6% accuracy) and the single layer perception variant, consisting of 500 neurons, discriminated environment types with 0.859 AUC (73.0% accuracy).
- Models trained on geographically distinct subgroups performed equally well when evaluated on the same data (p>0.05), suggesting good generalizability.
- The object detection framework, therefore, resulted in strong predictive performance similar in AUC over previous work, while achieving a 10-fold reduction in model complexity (i.e., number of model parameters) and providing new information about the composition of daily environments in which participants smoke.

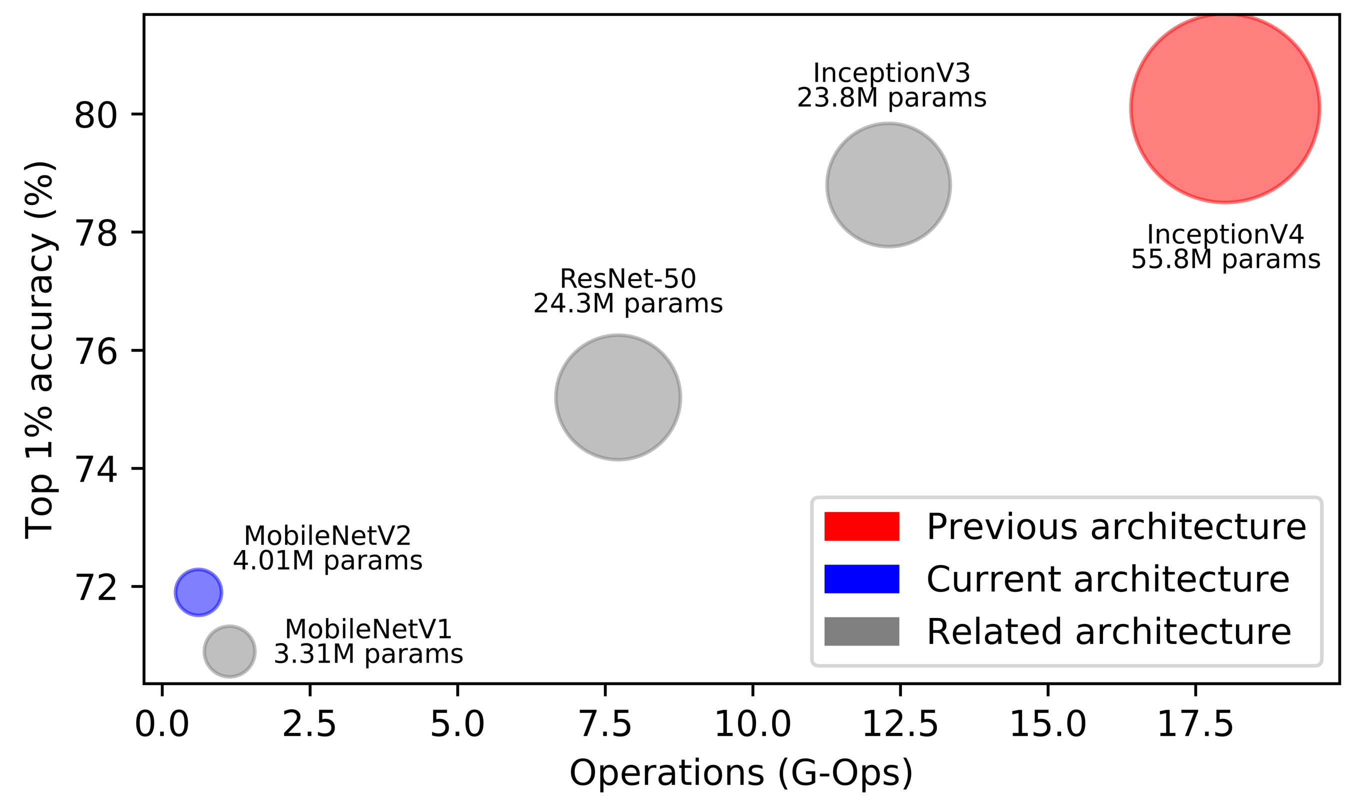

Figure 2: *A representation showing base architecture accuracy, number of operations, and number of parameters. Changing from a base architecture of InceptionV4 to MobileNetV2 caused a drop in parameters by 10-fold, while maintaining reasonable accuracy. Demonstrating pictorially why real-time smoking risk prediction is becoming a feasible.*

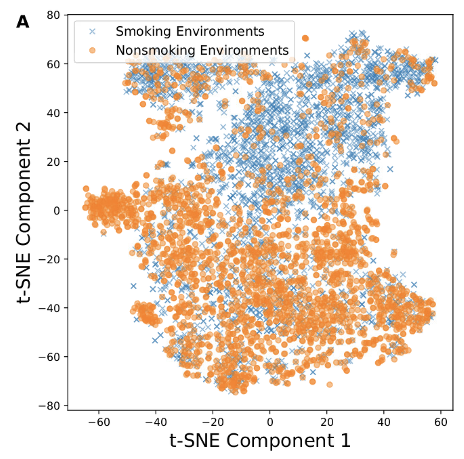

Figure 3: *A two-dimensional representation, created using t-stochastic neighbor embedding (t-SNE), of the image content extracted by the deep CNN shows how images taken by participants cluster into distinct environment types. Some are more likely to be smoking environments, and others are more likely to be nonsmoking environments.*

Acknowledgements: This work was funded by NIDA: R01 DA038442 and R21 DA033083 (McClernon); K23 DA042898 and R03 DA043756 (Oliver); R01 DA023646 (Conklin).

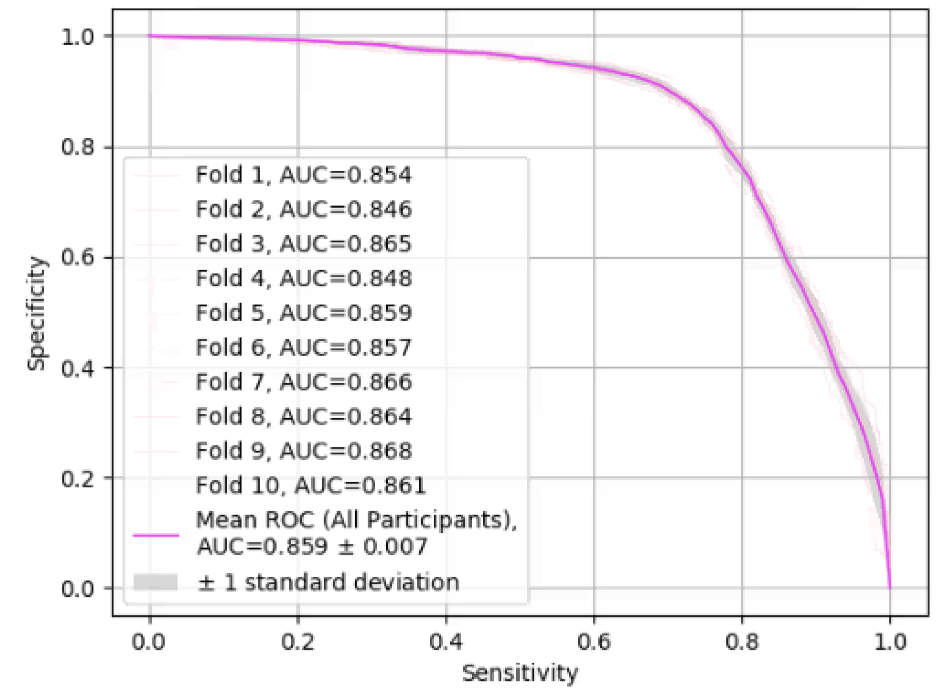

Figure 4: *Out-of-sample predictive performance (sensitivity vs specificity) for the trained classifier. The left panel shows mean ± SD of all results from cross-validation. Area under the curve (AUC) is 0.859 ± 0.007, with accuracy of 73.0% at an 0.5 classification threshold.*

Figure 5: *The combination of great model performance/generalizability due to the light object detection framework has made it feasible to pursue mobile app interventions. One possible idea is shown, where the app could predict risk for any area the user photographs, and draws bounding boxes around object proposals most correlated with smoking. Suggestions can also be made if the user is in an area of high smoking risk.*

4. Conclusions

- Object detection frameworks can **improve identification of smoking environments and predict smoking risk**:
  - → Identify and localize environmental features associated with smoking behavior
  - → Predict smoking risk associated with any image of daily life
  - → Predict risk in real time in order to trigger just-in-time, adaptive cessation interventions
- Good generalization across participants and geographic locations suggests that **specific environmental patterns are consistently associated with smoking**.
- Determining how external environments affect other behaviors or symptoms facilitates **environment-based interventions and therapeutic environment modifications**.