# Supplementary Material: Cross-Domain Forensic Shoeprint Matching

Bailey Kong[1]
bhkong@ics.uci.edu
James Supancic, III[1]
jsupanci@uci.edu
Deva Ramanan[2]
deva@andrew.cmu.edu
Charless Fowlkes[1]
fowlkes@ics.uci.edu

[1] Dept. of Computer Science
University of California, Irvine
United States of America
[2] Robotics Institute
Carnegie Mellon University
United States of America

## 1 Backpropagation for NCC

Here we derive $\frac{dNCC(x,y)}{dx}$. NCC is symmetric with respect to $x$ and $y$ so $\frac{dNCC(x,y)}{dy}$ is the same. NCC is a sum of terms over individual pixels $i$. Defining

$$NCC(x,y) = \frac{1}{|P|}\sum_i r_i(x,y)$$

we derive $\frac{dNCC(x,y)}{dx}$ by taking the total derivative:

$$\frac{dNCC(x,y)}{dx[j]} = \frac{1}{|P|}\sum_{i\in P}\frac{dr_i(x,y)}{dx[j]} \tag{1}$$

$$\frac{dr_i(x,y)}{dx[j]} = \tilde{y}[i]\frac{d\tilde{x}[i]}{dx[j]} \tag{2}$$

$$= \tilde{y}[i]\left(\frac{\partial\tilde{x}[i]}{\partial x[j]} + \frac{\partial\tilde{x}[i]}{\partial\mu_x}\frac{\partial\mu_x}{\partial x[j]} + \frac{\partial\tilde{x}[i]}{\partial\sigma_{xx}}\frac{\partial\sigma_{xx}}{\partial x[j]}\right) \tag{3}$$

The partial derivative $\frac{\partial\tilde{x}[i]}{\partial x[j]} = \frac{1}{\sqrt{\sigma_{xx}}}$, if and only if $i = j$ and is zero otherwise. The remaining partials derive as follows:

$$\frac{\partial\tilde{x}[i]}{\partial\mu_x} = -\frac{1}{\sqrt{\sigma_{xx}}} \qquad \frac{\partial\mu_x}{\partial x[j]} = \frac{1}{|P|} \tag{4}$$

$$\frac{\partial\tilde{x}[i]}{\partial\sigma_{xx}} = \frac{1}{2\sigma_{xx}^{3/2}}(x[i]-\mu_x) \qquad \frac{\partial\sigma_{xx}}{\partial x[j]} = \frac{2\,(x[j]-\mu_x)}{|P|} \tag{5}$$

Pulling everything together, we arrive at a final expression:

$$\frac{dNCC(x,y)}{dx[j]} = \frac{\tilde{y}[j]}{|P|\sqrt{\sigma_{xx}}} + \frac{1}{|P|}\sum_{i\in P}\tilde{y}[i]\left(\frac{-1}{|P|\sqrt{\sigma_{xx}}} + \frac{2\,(x[i]-\mu_x)\,(x[j]-\mu_x)}{2|P|\sigma_{xx}^{3/2}}\right) \quad (6)$$

$$= \frac{1}{|P|\sqrt{\sigma_{xx}}}\left(\tilde{y}[j] + \frac{1}{|P|}\sum_{i\in P}\tilde{y}[i]\left(-1 + \frac{(x[i]-\mu_x)\,(x[j]-\mu_x)}{\sigma_{xx}}\right)\right) \quad (7)$$

$$= \frac{1}{|P|\sqrt{\sigma_{xx}}}\left(\tilde{y}[j] - \frac{1}{|P|}\sum_{i\in P}\tilde{y}[i] + \frac{1}{|P|}\sum_{i\in P}\tilde{y}[i]\tilde{x}[i]\tilde{x}[j]\right) \quad (8)$$

$$= \frac{1}{|P|\sqrt{\sigma_{xx}}}(\tilde{y}[j] + \tilde{x}[j]NCC(x,y)) \quad (9)$$

where we have made use of the fact that $\tilde{y}$ is zero-mean.

# 2 Retrieval Results for Cross-Domain Matching

In this section we take a deeper look at our proposed system for the cross-domain matching problem, first looking at some qualitative results and second looking at its performance w.r.t. the size of crime scene prints.

In Fig. 2 and Fig. 3 we show the top-10 retrieved test impressions for a subset of crime scene prints from FID-300. These results correspond to $[\mu_c, \sigma_c]$ and $[\mu_c, \sigma_c \cdot W_c]$ of the right panel of Fig. 5 in the main paper.

Next we shows the performance of our two methods when evaluating crime scene prints of certain sizes. The size can give us an estimate of how much information is in the print—as crime scene prints and test impressions in FID-300 were scaled to a canonical size (10 pixels is 1 cm). For FID-300, the prints all fell into one of two categorizes which we call "full size" and "quarter size". "Full size" prints are crime scene prints whose pixel area is at least 90% of the corresponding test impression, whereas "quarter size" prints are crime scene prints whose pixel area is at most 25% of the corresponding test impression. 76 of the 300 crime scene prints were full size, while 224 of the 300 crime scene prints were quarter size.

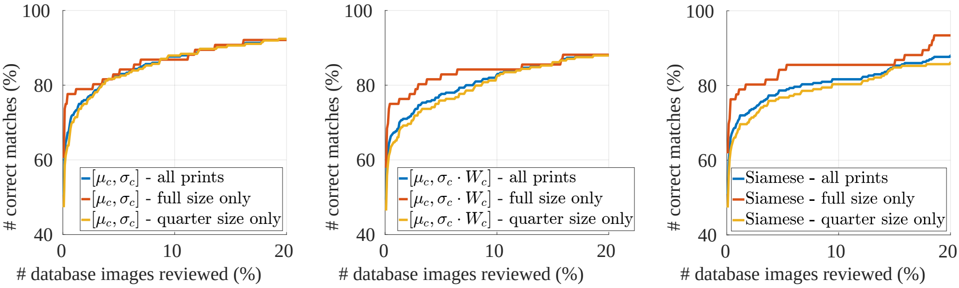


Figure 1: A study on the performance of our proposed methods on different sizes of crime scene prints in FID-300. The left panel shows the performance of MCNCC with uniform weights, the middle panel shows the performance of MCNCC with learned per-channel weights, and the right panel shows the performance of our Siamese network.

Fig. 1 shows us that, unsurprisingly, our system performs better on full size prints than on quarter size prints. This fits our intuition as larger prints will almost surely have more information to discriminate with. We also see our proposed system with learned per-channel weights performs worse than uniform weights on both categories. On the other hand, our Siamese network performs slightly better on full size prints but much worse on quarter size prints compared to uniform weights. These results reaffirm our suspicion that there is some overfitting in the models where we learn weights.

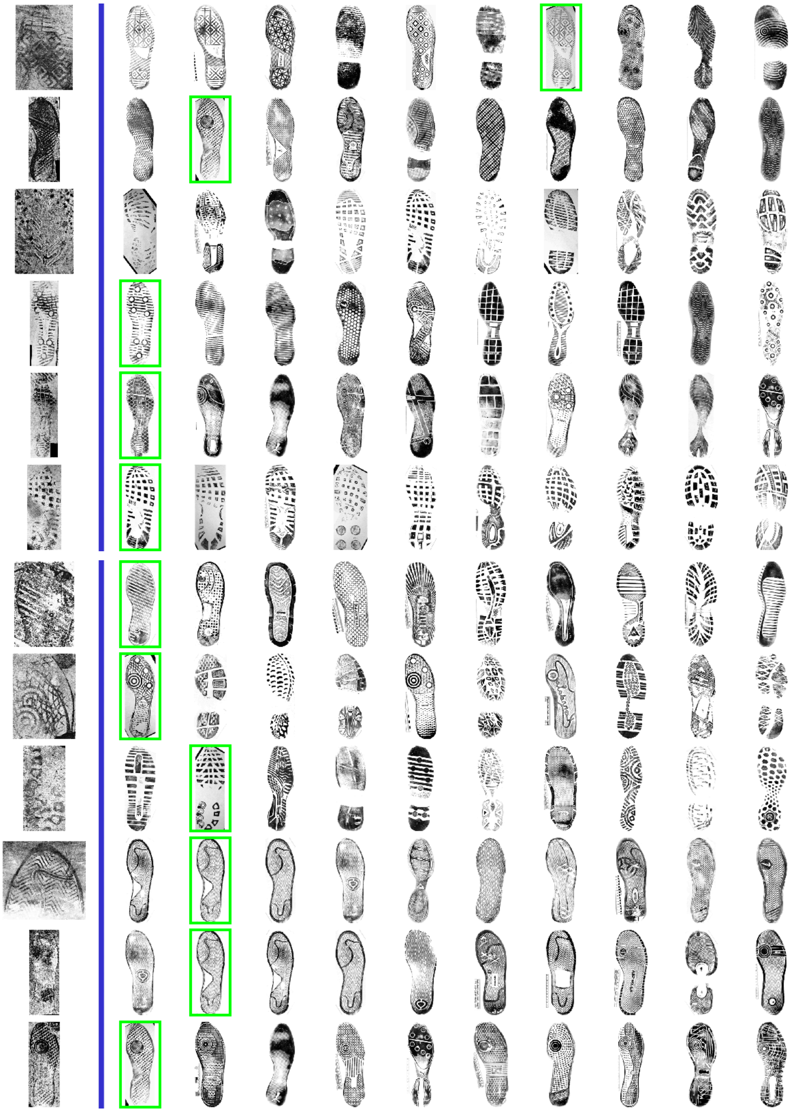

Figure 2: FID-300 retrieval results for $[\mu_c, \sigma_c]$. The left column shows the query crime scene prints. Green boxes indicate the corresponding ground truth test impression.

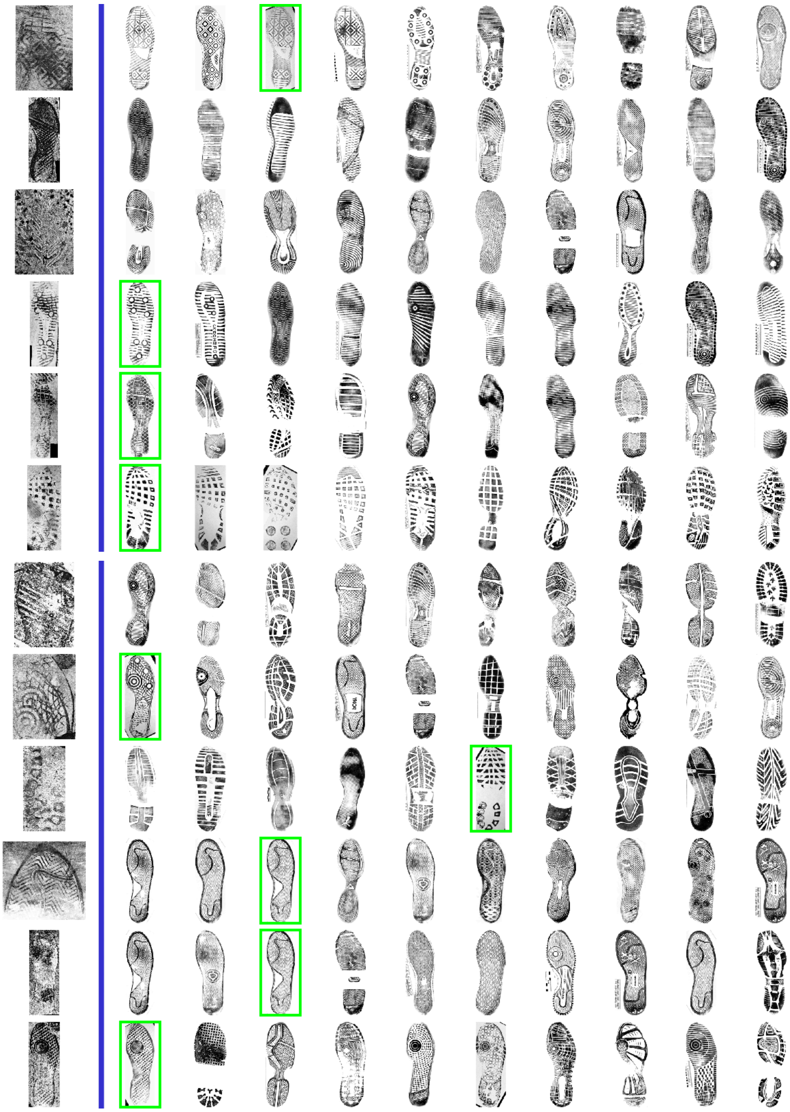

Figure 3: FID-300 retrieval results for $[\mu_c, \sigma_c \cdot W_c]$. The left column shows the query crime scene prints. Green boxes indicate the corresponding ground truth test impression.

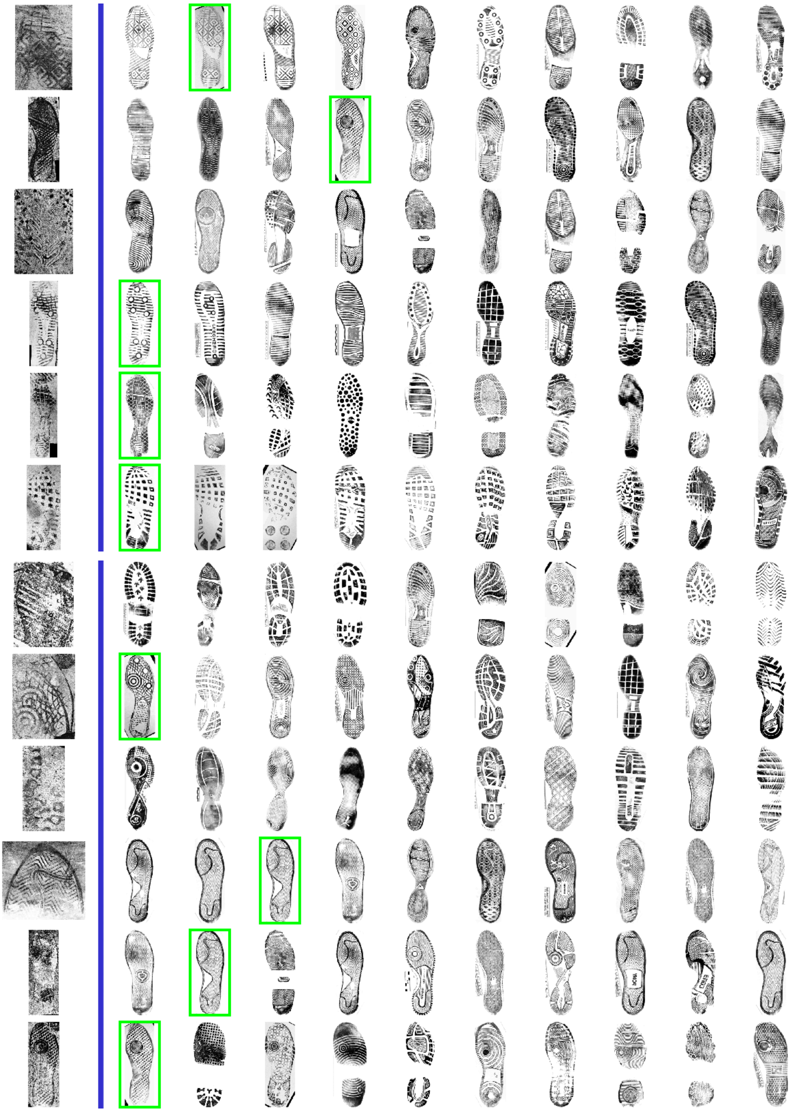

Figure 4: FID-300 retrieval results for "Siamese." The left column shows the query crime scene prints. Green boxes indicate the corresponding ground truth test impression.