Diversity and Novelty MasterPrints: Generating Multiple DeepMasterPrints for Increased User Coverage

M Charity, Nasir Memon, Zehua Jiang, Abhi Sen, Julian Togelius

Abstract

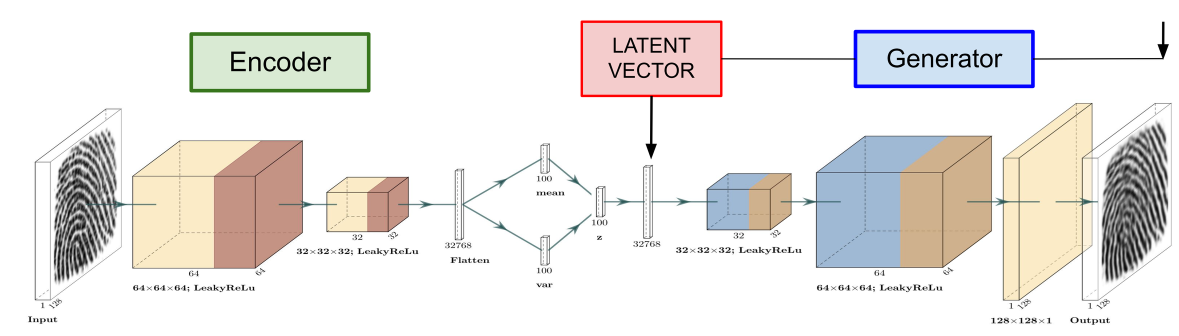
This work expands on previous advancements in genetic fingerprint spoofing via the DeepMasterPrints and introduces Diversity and Novelty MasterPrints. This system uses quality diversity evolutionary algorithms to generate dictionaries of artificial prints with a focus on increasing coverage of users from the dataset. The Diversity MasterPrints focus on generating solution prints that match with users not covered by previously found prints, and the Novelty MasterPrints explicitly search for prints with more that are farther in user space than previous prints. Our multi-print search methodologies outperform the singular DeepMasterPrints in both coverage and generalization while maintaining quality of the fingerprint image output.

Motivation

- Spoofing attacks focus a targeted individual
- Attackers would want to target as many individuals as possible
- Generating prints that can verify as multiple users with a single image
- Bontrager's work on Deep MasterPrints (DMP) generate such prints, however there are users missing from the coverage
- Creating more than one DMP does not focus on finding the remaining users and can have overlap with previously found users

Methods

- Generate an archive of Deep MasterPrints optimized for covering the remaining users in the user space for greater coverage
- Use a variational autoencoder (VAE) with covariance-matrix adaptation evolutionary strategy (CMA-ES) to create the diverse archive of prints
- The VAE creates more realistic prints from the latent space than the previously used GAN generator in Bontrager's work



Heuristics

- Diversity heuristic optimize in the CMA-ES function for remaining users not covered by prints in the archive
- Novelty heuristic optimize for prints that are farther away in the user space than prints in the archive
- Diversity is better for a known user space; Novelty is better for preventing overlap

Diversity heuristic

u_i = # unseen users from image i

U = # total users in coverage space

$\textbf{\textit{diversity\_value}}_i = u_i / U$

Novelty heuristic

$$novelty = \begin{cases} dist(x,0) & \text{if } len(d) = 0 \\ \min_{\forall s \in d} dist(x,s), & \text{otherwise} \end{cases}$$

Experiment Setup

- **Experiments**: (R)andom, (D)eep MasterPrints, D(I)versity Heuristic, (N)ovelty Heuristic
- **Dataset:** FingerPass DB7 128 x 128 grayscale
- **Verification**: Verifinger SDK
- **Generator**: Variational autoencoder
- **Archive size**: 10
- **Evolution Iterations**:
  10000 for Deep MasterPrints
  1000 for archive based prints
- **Trials**: 10
- Generated prints were trained on the first half of the real user fingerprints dataset and then evaluated on the remaining half of the dataset

Results

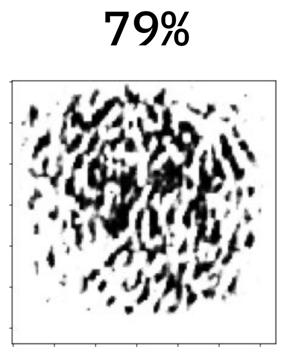
- The random prints generated with our VAE outperform Bontrager's WGAN randomly generated prints
- The match results from the Deep MasterPrints were on par with Bontrager's experiments
- Both heuristics outperform the Deep MasterPrints for coverage on both datasets
- Even at lower FMR rates, both prints managed to get almost double and triple the matches in the training set

VeriFinger Classifer		R	D	I	N
FMR 1.0	Train	52.66	78.83	**96.75**	95.97
	Test	53.77	72.72	93.66	**96.69**
FMR 0.1	Train	7.63	25.93	47.30	**48.25**
	Test	6.75	17.92	33.11	**40.86**
FMR 0.01	Train	0.58	3.63	**10.39**	10.19
	Test	0.14	0.95	**2.52**	2.38

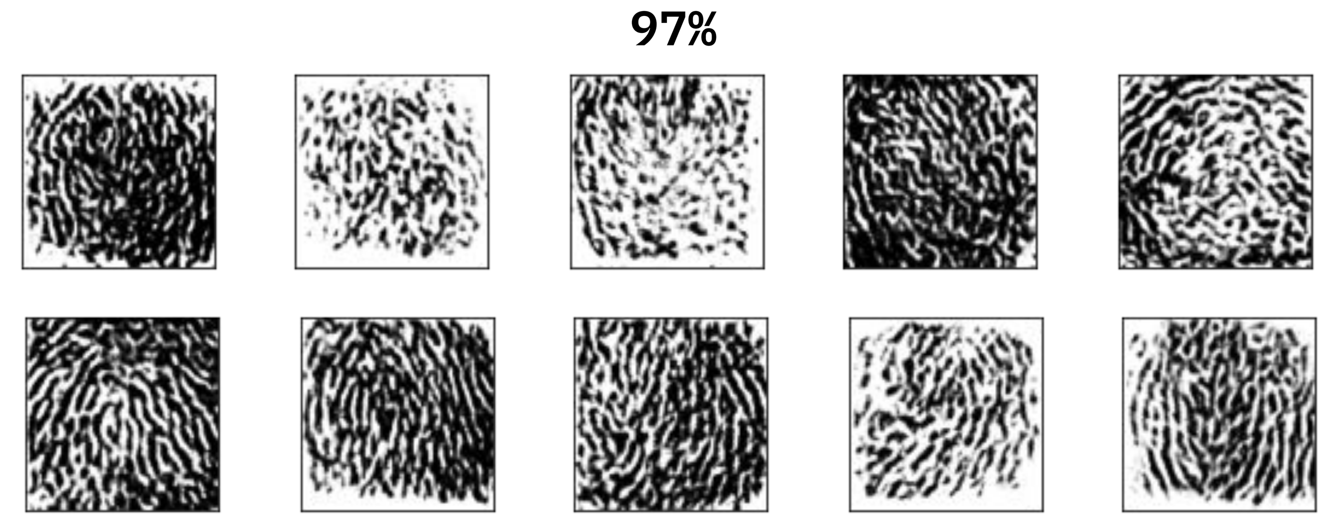
Print Output

Verifinger classifier - 1% FMR - Training dataset coverage - 360 users

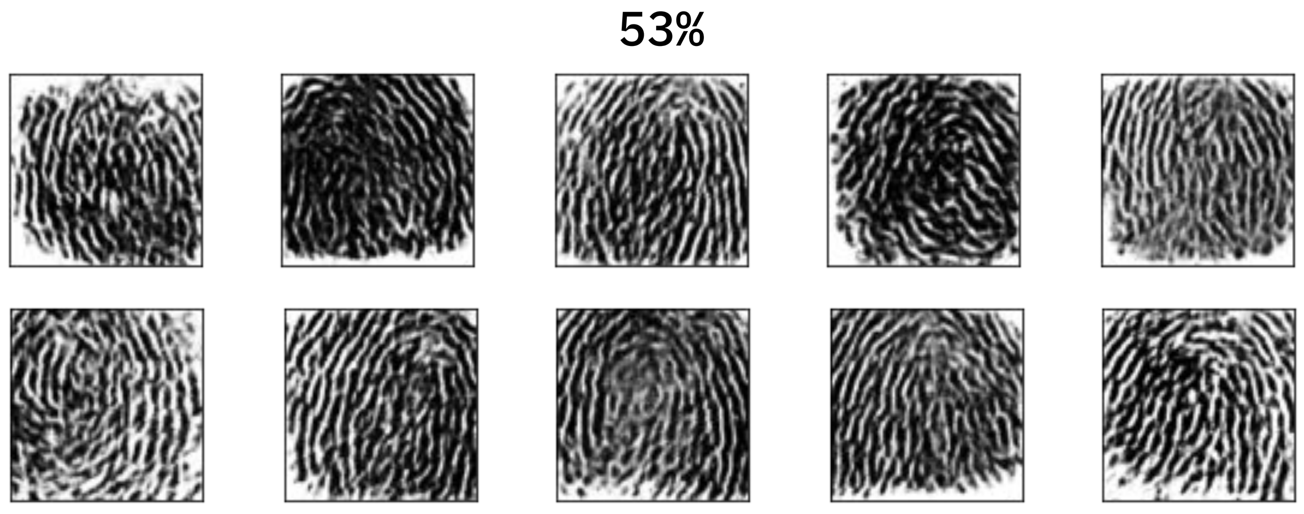
Deep MasterPrints
79%



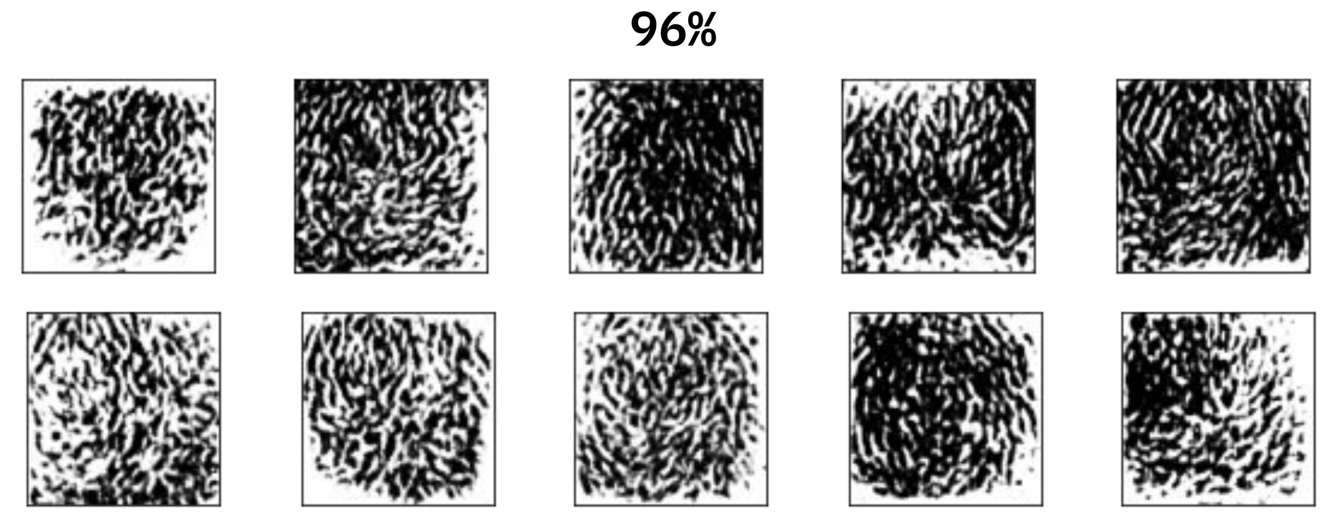
Diversity MasterPrints
97%



Randomly generated
53%



Novelty MasterPrints
96%



Conclusions and Future Work

- By creating a set of generated fingerprints with each new print attempting to cover the remaining subset of users, the match rate drastically increases over using a singular generated print that is evolved for the same amount of iterations, and also over multiple prints that are evolved independently.
- Apply the pipeline to other biometrics such as faces or voices
- Use other quality diversity algorithms such as MAP-Elites to find specific combinations of users
- Explore the variational autoencoder's latent vector to find more users quickly