CLAIMS

We Claim:

1. A computer-implemented method for improving security of a biometrics-based authentication system, comprising:

receiving, by one or more servers, enrolled biometric samples of an enrolled user during an enrollment stage of the biometrics-based authentication system;

creating, by the one or more servers, augmented biometric samples by adding learned perturbations to the enrolled biometric samples of the enrolled user;

receiving, by the one or more servers, during a request for authentication, submitted biometric samples from a second user purporting to be the enrolled user;

comparing, by the one or more servers, the submitted biometric samples of the second user to the enrolled biometric samples and to the augmented biometric samples of the enrolled user based on predefined metrics; and

based on the comparison, determining, by the one or more servers, that the submitted biometric samples of the second user have been modified to impersonate the enrolled user.

2. The method of claim 1, wherein receiving the enrolled biometric samples of the enrolled user during the enrollment stage, further comprises: applying a function $f(\cdot)$ to the enrolled biometric samples, x_i, to generate a biometric template $f(x_i)$.

3. The method of claim 2, further comprising: storing the enrolled biometric samples x_i, the biometric template $f(x_i)$ and the augmented biometric samples as an enrolled biometric dataset.

4. The method of claim 2, wherein receiving the submitted biometric samples from the second user, further comprises: applying the function $f(\cdot)$ to the submitted biometric

samples, x', to generate a biometric template f(x') to project the x_i and the x' to a common embedding subspace.

5. The method of claim 2, wherein comparing the submitted biometric samples of the second user to the enrolled biometric samples and to the augmented biometric samples of the enrolled user based on predefined metrics, further comprises:

computing a distance between the biometric template f(x_i) of the enrolled user with the biometric template f(x'); and

responsive to determining that the distance is less than a first threshold, temporarily authorizing the request for authentication.

6. The method of claim 5, wherein determining that the submitted biometric samples of the second user have been modified is activated only in response to the request for authentication being temporarily authorized.

7. The method of claim 6, further comprising:

responsive to the request for authentication being temporarily authorized, receiving, by an adversarial perturbation detector, the enrolled biometric samples x_i of the enrolled user and the submitted biometric samples of the second user, where the submitted biometric samples comprise perturbed adversarial samples x' + Δx'; and

responsive to detecting any perturbations, rejecting the request for authentication of the second user, and otherwise, granting the request for authentication and authorizing the second user.

8. The method of claim 7, wherein determining that any perturbations are detected further comprises:

applying a function k(·) to both the enrolled biometric samples x_i and the perturbed adversarial samples x' + Δx' to generate biometric template k(x_i) and a biometric template k(x' + Δx'), which are in a transformed subspace;

computing a distance F between biometric template $k(x_i)$ and the biometric template $k(x' + \Delta x')$ in the transformed subspace; and

determining the biometric template $k(x' + \sum \Delta x')$ is adversarial when the distance F is greater than a second threshold t' indicating that one or more perturbations have been detected.

9. The method of claim 8, wherein the enrolled biometric samples x_i include augmented biometric samples xi' such that the enrolled biometric samples xi = [xi, xi'], and wherein the x_i and x_i' of the enrolled biometric samples are input to the function $k(\cdot)$ using one of a parallel model and a sequential model.

10. The method of claim 8, further comprising:

generating a training enrolled biometric template $k(x_i)$, a training public biometric template k(x), and a training perturbed adversarial template $k(x' + \Delta x')$;

inputting the training examples to the function $k(\cdot)$, where $k(\cdot)$ has learnable parameters θ to generate a training enrolled biometric template $k(x_i)$, a training public biometric template k(x), and a training perturbed adversarial template $k(x' + \Delta x')$;

classifying, by a classifier with learnable parameters σ, the training perturbed adversarial template $k(x' + \Delta x')$ as a 1 and the training public biometric template k(x) as 0 based on the training enrolled biometric template $k(x_i)$; and

based on a result of the classification, updating the learnable parameters θ and σ.

11. A non-transitory computer readable medium having stored thereon software instructions that, when executed by a processor, cause the processor to improve security of a biometrics-based authentication system, the instructions comprising:

receiving, by one or more servers, enrolled biometric samples of an enrolled user during an enrollment stage of the biometrics-based authentication system;

creating, by the one or more servers, augmented biometric samples by adding learned perturbations to the enrolled biometric samples of the enrolled user;

receiving, by the one or more servers, during a request for authentication, submitted biometric samples from a second user purporting to be the enrolled user;

comparing, by the one or more servers, the submitted biometric samples of the second user to the enrolled biometric samples and to the augmented biometric samples of the enrolled user based on predefined metrics; and

based on the comparison, determining, by the one or more servers, that the submitted biometric samples of the second user have been modified to impersonate the enrolled user.

12. The non-transitory computer readable medium of claim 11, wherein receiving the enrolled biometric samples of the enrolled user during the enrollment stage, further comprises: applying a function $f(\cdot)$ to the enrolled biometric samples, x_i, to generate a biometric template $f(x_i)$.

13. The non-transitory computer readable medium of claim 12, further comprising: storing the enrolled biometric samples x_i, the biometric template $f(x_i)$ and the augmented biometric samples as an enrolled biometric dataset.

14. The non-transitory computer readable medium of claim 12, wherein receiving the submitted biometric samples from the second user, further comprises: applying the function $f(\cdot)$ to the submitted biometric samples, x', to generate a biometric template f(x') to project the x_i and the x' to a common embedding subspace.

15. The non-transitory computer readable medium of claim 12, wherein comparing the submitted biometric samples of the second user to the enrolled biometric samples and to the augmented biometric samples of the enrolled user based on predefined metrics, further comprises:

computing a distance between the biometric template $f(x_i)$ of the enrolled user with the biometric template f(x'); and

responsive to determining that the distance is less than a first threshold, temporarily authorizing the request for authentication.

16. The non-transitory computer readable medium of claim 15, wherein determining that the submitted biometric samples of the second user have been modified is activated only in response to the request for authentication being temporarily authorized.

17. The non-transitory computer readable medium of claim 16, further comprising:

responsive to the request for authentication being temporarily authorized, receiving, by an adversarial perturbation detector, the enrolled biometric samples x_i of the enrolled user and the submitted biometric samples of the second user, where the submitted biometric samples comprise perturbed adversarial samples $x' + \Delta x'$; and

responsive to detecting any perturbations, rejecting the request for authentication of the second user, and otherwise, granting the request for authentication and authorizing the second user.

18. The non-transitory computer readable medium of claim 17, wherein determining that any perturbations are detected further comprises:

applying a function $k(\cdot)$ to both the enrolled biometric samples x_i and the perturbed adversarial samples $x' + \Delta x'$ to generate biometric template $k(x_i)$ and a biometric template $k(x' + \Delta x')$, which are in a transformed subspace;

computing a distance F between biometric template $k(x_i)$ and the biometric template $k(x' + \Delta x')$ in the transformed subspace; and

determining the biometric template $k(x' + \sum \Delta x')$ is adversarial when the distance F is greater than a second threshold t' indicating that one or more perturbations have been detected.

19. The non-transitory computer readable medium of claim 18, further comprising:

generating a training enrolled biometric template $k(x_i)$, a training public biometric template k(x), and a training perturbed adversarial template $k(x' + \Delta x')$;

inputting the training examples to the function $k(\cdot)$, where $k(\cdot)$ has learnable parameters θ to generate a training enrolled biometric template $k(x_i)$, a training public biometric template k(x), and a training perturbed adversarial template $k(x' + \Delta x')$;

classifying, by a classifier with learnable parameters σ, the training perturbed adversarial template $k(x' + \Delta x')$ as a 1 and the training public biometric template k(x) as 0 based on the training enrolled biometric template $k(x_i)$; and

based on a result of the classification, updating the learnable parameters θ and σ.

20. A system, comprising:

a memory;

a processor coupled to the memory; and

a software component executed by the processor that is configured to:

receive enrolled biometric samples of an enrolled user during an enrollment stage of the biometrics-based authentication system;

create augmented biometric samples by adding learned perturbations to the enrolled biometric samples of the enrolled user;

receive during a request for authentication, submitted biometric samples from a second user purporting to be the enrolled user;

compare the submitted biometric samples of the second user to the enrolled biometric samples and to the augmented biometric samples of the enrolled user based on predefined metrics; and

based on the comparison, determine that the submitted biometric samples of the second user have been modified to impersonate the enrolled user.