

End User License Agreement

(Academic, non-commercial, not-for-profit licence)

Copyright (c) 2025 Kartik Kuckreja, Parul Gupta, Injy Hamed, Thamar Solorio, Muhammad Haris Khan, and Abhinav Dhall All rights reserved.

The goal of the ArEnAV database is to develop new techniques, technology, and algorithms for multimodal, code-switched deepfake detection and localization, as most of the existing research focuses on monolingual content, often overlooking the challenges of multilingual and code-switched speech, where multiple languages are mixed within the same discourse. The licensors are involved in an ongoing effort to strengthen detection algorithms against highly realistic deepfakes. The dataset is meant to aid research efforts in the general area of developing, testing and evaluating algorithms for multilingual code-switched deepfake detection and localization.

To receive a copy of the dataset, the requestor must agree to observe the conditions listed Below.

The goal of the ArEnAV database is to develop new techniques, technology, and algorithms for predicting and locating (with timestamps) where a video has been manipulated, particularly when it has Arabic-English code-switching. Use is permitted of the databases and annotations above in source and binary form, provided that the following conditions are met:

- The database is provided under the terms of this license strictly for academic, non-commercial, not-for-profit purposes.
- Redistribution, republishing, or dissemination in any form, source or binary, is not permitted without prior written approval by the licensors. Linking to the webpage of the database [INSERT WEB LINK HERE] is permitted.
- The names of the licensors may not be used to endorse or promote products derived from this software without specific prior written permission.
- The licensors reserve the right to modify the data/license at any point. Modification of the database by licensees is not permitted.
- In no case should the still frames or videos be used in any way that could cause the original subject embarrassment or mental anguish.
- You understand that the ArEnAV dataset is a deepfake dataset generated based on VisPer ([2406.00038] ViSpeR: Multilingual Audio-Visual Speech Recognition) dataset's Arabic Train subset. You also agree to all agreements of the VisPer dataset.
- The authors of the dataset make no representations or warranties regarding the dataset, including but not limited to warranties of non-infringement or fitness for a particular purpose.
- You accept full responsibility for your use of the dataset and shall defend and indemnify the Authors of ArEnEV, against any and all claims arising from your use of the dataset, including but not limited to your use of any copies of copyrighted images that you may create from the dataset.

- Any publications arising from the use of this software, including but not limited to academic journal and conference publications, technical reports and manuals, must cite the following works:

```
@misc{kuckreja2025tellhabibirealfake,
  title={Tell me Habibi, is it Real or Fake?},
  author={Kartik Kuckreja and Parul Gupta and Injy Hamed and Tamar Solorio and
Muhammad Haris Khan and Abhinav Dhall},
  year={2025},
  eprint={2505.22581},
  archivePrefix={arXiv},
  primaryClass={cs.CV},
  url={https://arxiv.org/abs/2505.22581},
}
```

THE DATABASE IS PROVIDED BY THE AUTHORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS DATABASE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE. THE PROVIDER OF THE DATABASE MAKES NO REPRESENTATIONS AND EXTENDS NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED. THERE ARE NO EXPRESS OR IMPLIED WARRANTIES THAT THE USE OF THE MATERIAL WILL NOT INFRINGE ANY PATENT, COPYRIGHT, TRADEMARK, OR OTHER PROPRIETARY RIGHTS.

If you have read and understood the user agreement and will comply with it.

Signed _____

Print Name _____

Institution Name _____

Date _____

Addition Researcher 1 _____

Addition Researcher 2 _____

Addition Researcher 3 _____

Addition Researcher 4 _____

Addition Researcher 5 _____

Addition Researcher 6 _____

Addition Researcher 7 _____