**King Fahd University Arabic Font Database (KAFD)**

The database contains printed Arabic text images and its ground-truth developed for research in the area of Arabic text recognition. KAFD is a multi-font, multi-size, multi-style, and multi-resolution Arabic text database.

**Overview:**

* Arabic texts are collected from different subjects like religious, medicine, science, history…etc. The used texts cover all the shapes of Arabic characters.
* The collected texts are printed using (40) fonts listed in Table 1.
* Arabic text is printed using HP Laser jet 600 M601 printer with a print resolution of 1200 x 1200 dpi.
* The printed pages are scanned in grayscale using Ricoh IS760D scanner using 100 dpi, 200 dpi, 300 dpi and 600 dpi resolutions.
* Arabic texts include ten sizes (8, 9, 10, 11, 12, 14, 16, 18, 20, and 24 points).
* For each size, four font styles are used (Normal, Bold, Italic, and a combination of Bold and Italic).
* For each font style, three disjoint sets are constructed (60% for Training, 20% for Testing, and 20% Validation).
* The scanning process resulted in **(115,068)** page level text images for all resolutions (28,767 page images per resolution)
* All text images are segmented into lines. This process resulted in **(2,576,024)** line level text images (644,006 line images per resolution).
* Page and line images are supplied with manually verified ground-truths.
* KAFD is available in two forms: page and line text images. The page level of KAFD database consists of (28,767) text images per resolution. This number of page text images is distributed between Training (17,131), Testing (5,816), and Validation (5,820) sets. A total of (115,068) page images are the number of page images of KAFD database as shown in Table 2.
* The line level of KAFD database consists of (644,006) text images per resolution. This number is distributed between Training (389,736), Testing (126,323), and Validation (127,947) sets. A total of (2,576,024) line images are the number of line images of K-AFD database as shown in Table 2.



**Table 1. KAFD fonts**

**Table 2. KAFD statistics**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.N** | **Font** | **Page level database** | **Line level database** |
| **Training set** | **Testing set** | **Validation set** | **Total** | **Total** | **Training set** | **Testing set** | **Validation set** | **Total** | **Total** |
| 1 | Advertising Bold | 430 | 152 | 144 | 726 | 2904 | 9718 | 3342 | 2982 | 16042 | 64168 |
| 2 | AGA Granada Regular | 425 | 144 | 136 | 705 | 2820 | 9265 | 3196 | 2738 | 15199 | 60796 |
| 3 | AGA Kaleelah Regular | 440 | 152 | 150 | 742 | 2968 | 9958 | 3244 | 3218 | 16420 | 65680 |
| 4 | Akhbar | 431 | 147 | 150 | 728 | 2912 | 8432 | 2715 | 2982 | 14129 | 56516 |
| 5 | Al-Mohand | 425 | 144 | 152 | 721 | 2884 | 9134 | 2813 | 3161 | 15108 | 60432 |
| 6 | AL-Qairwan | 428 | 152 | 144 | 724 | 2896 | 9152 | 3135 | 2928 | 15215 | 60860 |
| 7 | Andalus | 424 | 145 | 144 | 713 | 2852 | 8806 | 2920 | 2968 | 14694 | 58776 |
| 8 | Arabic Transparent | 430 | 150 | 148 | 728 | 2912 | 12591 | 4228 | 4176 | 20995 | 83980 |
| 9 | Arabic Typesetting | 428 | 144 | 144 | 716 | 2864 | 11482 | 3832 | 3676 | 18990 | 75960 |
| 10 | Arabswell | 434 | 148 | 148 | 730 | 2920 | 7948 | 2578 | 2582 | 13108 | 52432 |
| 11 | Arial | 419 | 140 | 144 | 703 | 2812 | 11181 | 3560 | 3756 | 18497 | 73988 |
| 12 | Arial Unicode MS | 438 | 145 | 152 | 735 | 2940 | 10847 | 3534 | 3876 | 18257 | 73028 |
| 13 | Courier New | 412 | 144 | 148 | 704 | 2816 | 11499 | 3868 | 3997 | 19364 | 77456 |
| 14 | Deco Type Naskh | 437 | 150 | 149 | 736 | 2944 | 6960 | 2230 | 2302 | 11492 | 45968 |
| 15 | Deco Type Thuluth | 428 | 144 | 152 | 724 | 2896 | 6950 | 2218 | 2392 | 11560 | 46240 |
| 16 | Diwani Letter | 428 | 148 | 144 | 720 | 2880 | 7078 | 2388 | 2208 | 11674 | 46696 |
| 17 | Freehand | 432 | 148 | 148 | 728 | 2912 | 9554 | 3136 | 3094 | 15784 | 63136 |
| 18 | Hadeel | 436 | 132 | 136 | 704 | 2816 | 9650 | 2542 | 2816 | 15008 | 60032 |
| 19 | Lotus Linotype  | 436 | 148 | 152 | 736 | 2944 | 7580 | 2392 | 2568 | 12540 | 50160 |
| 20 | M Unicode Sara | 434 | 144 | 152 | 730 | 2920 | 9534 | 3024 | 3300 | 15858 | 63432 |
| 21 | MaghribiAssile | 432 | 151 | 147 | 730 | 2920 | 12209 | 4079 | 3785 | 20073 | 80292 |
| 22 | Microsoft Sans Serif | 412 | 145 | 136 | 693 | 2772 | 11868 | 4020 | 3784 | 19672 | 78688 |
| 23 | Microsoft Uighur | 447 | 140 | 139 | 726 | 2904 | 13613 | 4028 | 5046 | 22687 | 90748 |
| 24 | Midan | 428 | 138 | 144 | 710 | 2840 | 11930 | 3608 | 3836 | 19374 | 77496 |
| 25 | Motken Unicode Hor | 428 | 144 | 152 | 724 | 2896 | 9208 | 2976 | 3131 | 15315 | 61260 |
| 26 | Nawel | 416 | 148 | 144 | 708 | 2832 | 9038 | 3104 | 2982 | 15124 | 60496 |
| 27 | Pen Kufi | 423 | 148 | 149 | 720 | 2880 | 8129 | 2818 | 2836 | 13783 | 55132 |
| 28 | Quran2 | 425 | 144 | 148 | 717 | 2868 | 6900 | 2140 | 2262 | 11302 | 45208 |
| 29 | Rateb | 422 | 147 | 148 | 717 | 2868 | 8646 | 2956 | 2968 | 14570 | 58280 |
| 30 | Rekaa | 424 | 136 | 146 | 706 | 2824 | 7296 | 2180 | 2484 | 11960 | 47840 |
| 31 | SC Dubai | 432 | 152 | 152 | 736 | 2944 | 10900 | 3620 | 3708 | 18228 | 72912 |
| 32 | SC Gulf | 436 | 144 | 132 | 712 | 2848 | 7922 | 2456 | 2186 | 12564 | 50256 |
| 33 | Sc-alyermook | 424 | 136 | 136 | 696 | 2784 | 10900 | 3516 | 3547 | 17963 | 71852 |
| 34 | sc-ouhod | 416 | 148 | 146 | 710 | 2840 | 15344 | 5086 | 4946 | 25376 | 101504 |
| 35 | Segore UI | 419 | 142 | 147 | 708 | 2832 | 10117 | 3320 | 3474 | 16911 | 67644 |
| 36 | Simplified Arabic | 430 | 154 | 154 | 738 | 2952 | 8213 | 2772 | 2768 | 13753 | 55012 |
| 37 | Tahoma | 436 | 150 | 150 | 736 | 2944 | 11594 | 3852 | 3803 | 19249 | 76996 |
| 38 | Times New Roman | 430 | 144 | 144 | 718 | 2872 | 11563 | 3679 | 3714 | 18956 | 75824 |
| 39 | Traditional Arabic | 432 | 149 | 140 | 721 | 2884 | 8827 | 2928 | 2716 | 14471 | 57884 |
| 40 | Zarnew | 424 | 135 | 129 | 688 | 2752 | 8200 | 2290 | 2251 | 12741 | 50964 |
| **Total** | **17,131** | **5,816** | **5,820** | **28,767** | **115,068** | **389,736** | **126,323** | **127,947** | **644,006** | **2,576,024** |

**Reference**

1. *Luqman, Hamzah, Sabri A. Mahmoud, and Sameh Awaida.* "KAFD Arabic Font Database." *Pattern Recognition (2014).‏*