**Supplementary material of the paper**

**Jovan Galić, Branko Marković, Đorđe Grozdić, Branislav Popović and Slavko Šajić  
“Whispered Speech Recognition Based on Audio Data Augmentation and Inverse Filtering”**



Figure 9. The average recognition accuracy (in %) for Whi-Spe (closed set) and DBtest database (open set) in the HMM framework. The horizontal axis denotes the percentage of the Whi-Spe subset employed in training.



Figure 10. The average recognition accuracy (in %) for Whi-Spe (closed set) and DBtest database (open set) in the CNN framework. The horizontal axis denotes the percentage of the Whi-Spe subset employed in training.

HMM (Whi-Spe) – Exp\_1

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 10% | 94.4 | 95 | 91.2 | 94.8 | 86.4 | 96.6 | 95 | 95.6 | 92.6 | 95.2 |
| 20% | 94.8 | 95.2 | 92.6 | 96 | 87 | 97.2 | 95.8 | 96.2 | 92.2 | 94.4 |
| 30% | 95.6 | 94.6 | 94.4 | 96 | 88.4 | 97 | 96 | 95.8 | 92.4 | 95 |
| 40% | 95.8 | 95.4 | 94.2 | 96.8 | 88.8 | 96.8 | 96.6 | 95.6 | 93.8 | 95 |
| 50% | 95.8 | 95.8 | 94 | 96.8 | 88.8 | 97 | 96.4 | 95.2 | 94 | 94.8 |
| 60% | 95.2 | 96.6 | 93.4 | 96 | 89.2 | 96.6 | 96.4 | 95 | 93.4 | 93.6 |
| 70% | 95.6 | 97 | 93.8 | 96.2 | 89.6 | 97.4 | 96.8 | 95.6 | 93.2 | 94.8 |
| 80% | 95.4 | 97.0 | 93.8 | 96 | 89 | 97.2 | 96.6 | 95.8 | 93.6 | 94 |
| 90% | 95.6 | 97.2 | 93.6 | 96.4 | 89.6 | 96.8 | 96.2 | 96 | 94 | 94.4 |
| 100% | 95.4 | 97.2 | 93.6 | 96.2 | 89.4 | 97.2 | 96.2 | 96.2 | 94.4 | 94.4 |

CNN (Whi-Spe) – Exp\_2

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 10% | 84.6 | 83.6 | 81 | 85.2 | 74.8 | 69.6 | 80.8 | 72.2 | 60.2 | 68 |
| 20% | 90.8 | 90.8 | 83.8 | 85.8 | 84 | 82.4 | 87.2 | 86.2 | 69.8 | 75.8 |
| 30% | 92.2 | 92.4 | 89.4 | 93.2 | 88.6 | 87 | 91 | 90.6 | 74.8 | 78 |
| 40% | 93.2 | 92.6 | 92 | 95.2 | 89.2 | 88.8 | 93.4 | 88.4 | 73.2 | 79.2 |
| 50% | 94.6 | 93.8 | 94.4 | 96.8 | 90.4 | 88.2 | 92.8 | 92.4 | 77.2 | 82.2 |
| 60% | 94.8 | 95.4 | 96 | 96.2 | 92 | 90 | 95.4 | 92.8 | 78.6 | 80.6 |
| 70% | 95.4 | 93.2 | 96.4 | 94.2 | 94.2 | 91.6 | 96 | 95.2 | 78 | 84 |
| 80% | 94 | 95.6 | 96.4 | 95.6 | 92.6 | 90.8 | 96.2 | 94 | 82.4 | 85.6 |
| 90% | 95.2 | 95.8 | 97 | 97 | 94.6 | 89.2 | 96 | 92.6 | 79.2 | 86.6 |
| 100% | 94.6 | 96.2 | 97.4 | 97.2 | 93.4 | 90.8 | 96.2 | 93.2 | 81 | 84.6 |

CNN (DBtest)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 100% | 0.88 | 0.65 | 0.85 | 0.93 | 0.92 | 0.77 | 0.77 | 0.92 | 0.91 | 0.73 |
| 0.9 | 0.64 | 0.87 | 0.95 | 0.9 | 0.87 | 0.76 | 0.94 | 0.88 | 0.77 |
| 0.92 | 0.57 | 0.91 | 0.95 | 0.94 | 0.79 | 0.83 | 0.93 | 0.93 | 0.8 |
| 0.92 | 0.59 | 0.82 | 0.93 | 0.94 | 0.79 | 0.79 | 0.93 | 0.87 | 0.74 |
| 0.88 | 0.65 | 0.88 | 0.98 | 0.94 | 0.87 | 0.8 | 0.91 | 0.89 | 0.73 |
| 0.91 | 0.61 | 0.85 | 0.93 | 0.95 | 0.8 | 0.8 | 0.92 | 0.88 | 0.76 |
| 0.88 | 0.64 | 0.87 | 0.94 | 0.89 | 0.78 | 0.82 | 0.9 | 0.82 | 0.72 |
| 0.86 | 0.66 | 0.89 | 0.94 | 0.89 | 0.77 | 0.79 | 0.9 | 0.87 | 0.78 |
| 0.85 | 0.65 | 0.87 | 0.95 | 0.93 | 0.77 | 0.81 | 0.91 | 0.87 | 0.78 |
| 0.87 | 0.66 | 0.91 | 0.96 | 0.92 | 0.8 | 0.76 | 0.93 | 0.94 | 0.75 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 90% | 0.88 | 0.65 | 0.87 | 0.95 | 0.93 | 0.77 | 0.8 | 0.92 | 0.86 | 0.74 |
| 0.92 | 0.61 | 0.87 | 0.94 | 0.96 | 0.77 | 0.79 | 0.87 | 0.9 | 0.75 |
| 0.87 | 0.61 | 0.92 | 0.95 | 0.94 | 0.83 | 0.8 | 0.92 | 0.87 | 0.73 |
| 0.89 | 0.68 | 0.91 | 0.97 | 0.96 | 0.81 | 0.83 | 0.93 | 0.92 | 0.76 |
| 0.9 | 0.58 | 0.87 | 0.97 | 0.95 | 0.79 | 0.82 | 0.92 | 0.89 | 0.74 |
| 0.91 | 0.59 | 0.84 | 0.96 | 0.91 | 0.75 | 0.74 | 0.9 | 0.88 | 0.71 |
| 0.86 | 0.59 | 0.86 | 0.92 | 0.9 | 0.79 | 0.82 | 0.95 | 0.88 | 0.71 |
| 0.9 | 0.64 | 0.9 | 0.99 | 0.91 | 0.76 | 0.77 | 0.89 | 0.87 | 0.78 |
| 0.85 | 0.63 | 0.85 | 0.96 | 0.93 | 0.83 | 0.82 | 0.9 | 0.89 | 0.76 |
| 0.87 | 0.68 | 0.85 | 0.95 | 0.94 | 0.78 | 0.75 | 0.89 | 0.83 | 0.71 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 80% | 0.86 | 0.62 | 0.87 | 0.96 | 0.89 | 0.85 | 0.81 | 0.92 | 0.86 | 0.73 |
| 0.86 | 0.59 | 0.88 | 0.95 | 0.92 | 0.79 | 0.77 | 0.93 | 0.83 | 0.65 |
| 0.85 | 0.64 | 0.86 | 0.97 | 0.9 | 0.74 | 0.76 | 0.89 | 0.86 | 0.71 |
| 0.92 | 0.6 | 0.9 | 0.97 | 0.91 | 0.83 | 0.76 | 0.92 | 0.87 | 0.71 |
| 0.91 | 0.69 | 0.81 | 0.97 | 0.92 | 0.82 | 0.83 | 0.92 | 0.93 | 0.77 |
| 0.9 | 0.57 | 0.88 | 0.96 | 0.92 | 0.76 | 0.81 | 0.91 | 0.84 | 0.77 |
| 0.9 | 0.62 | 0.86 | 0.92 | 0.91 | 0.81 | 0.77 | 0.92 | 0.88 | 0.73 |
| 0.91 | 0.64 | 0.85 | 0.97 | 0.86 | 0.81 | 0.8 | 0.92 | 0.79 | 0.74 |
| 0.83 | 0.59 | 0.87 | 0.96 | 0.91 | 0.77 | 0.8 | 0.96 | 0.86 | 0.7 |
| 0.92 | 0.63 | 0.84 | 0.96 | 0.9 | 0.78 | 0.79 | 0.89 | 0.9 | 0.74 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 70% | 0.83 | 0.62 | 0.84 | 0.96 | 0.87 | 0.78 | 0.78 | 0.87 | 0.77 | 0.76 |
| 0.8 | 0.68 | 0.85 | 0.95 | 0.95 | 0.76 | 0.73 | 0.9 | 0.86 | 0.74 |
| 0.88 | 0.65 | 0.89 | 0.94 | 0.88 | 0.75 | 0.81 | 0.89 | 0.91 | 0.7 |
| 0.87 | 0.59 | 0.88 | 0.96 | 0.9 | 0.82 | 0.8 | 0.9 | 0.84 | 0.74 |
| 0.91 | 0.61 | 0.86 | 0.94 | 0.93 | 0.78 | 0.8 | 0.88 | 0.89 | 0.78 |
| 0.85 | 0.63 | 0.87 | 0.98 | 0.91 | 0.75 | 0.79 | 0.9 | 0.9 | 0.71 |
| 0.89 | 0.64 | 0.78 | 0.98 | 0.9 | 0.77 | 0.74 | 0.87 | 0.86 | 0.67 |
| 0.86 | 0.6 | 0.87 | 0.96 | 0.94 | 0.83 | 0.76 | 0.93 | 0.91 | 0.77 |
| 0.88 | 0.6 | 0.87 | 0.96 | 0.91 | 0.83 | 0.78 | 0.89 | 0.9 | 0.77 |
| 0.9 | 0.63 | 0.85 | 0.96 | 0.92 | 0.83 | 0.85 | 0.88 | 0.91 | 0.73 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 60% | 0.87 | 0.57 | 0.84 | 0.97 | 0.89 | 0.73 | 0.82 | 0.91 | 0.85 | 0.67 |
| 0.78 | 0.58 | 0.8 | 0.94 | 0.91 | 0.75 | 0.69 | 0.88 | 0.78 | 0.67 |
| 0.85 | 0.55 | 0.87 | 0.96 | 0.89 | 0.77 | 0.7 | 0.92 | 0.85 | 0.7 |
| 0.86 | 0.62 | 0.83 | 0.98 | 0.9 | 0.82 | 0.74 | 0.91 | 0.88 | 0.7 |
| 0.86 | 0.64 | 0.87 | 0.96 | 0.91 | 0.78 | 0.71 | 0.91 | 0.84 | 0.77 |
| 0.86 | 0.58 | 0.85 | 0.94 | 0.9 | 0.76 | 0.75 | 0.91 | 0.79 | 0.7 |
| 0.85 | 0.61 | 0.85 | 0.96 | 0.9 | 0.78 | 0.79 | 0.89 | 0.86 | 0.74 |
| 0.91 | 0.62 | 0.88 | 0.96 | 0.96 | 0.8 | 0.77 | 0.91 | 0.88 | 0.74 |
| 0.83 | 0.64 | 0.8 | 0.94 | 0.86 | 0.8 | 0.63 | 0.9 | 0.84 | 0.69 |
| 0.84 | 0.62 | 0.89 | 0.94 | 0.9 | 0.76 | 0.81 | 0.91 | 0.92 | 0.76 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 50% | 0.87 | 0.56 | 0.85 | 0.88 | 0.91 | 0.72 | 0.7 | 0.88 | 0.81 | 0.75 |
| 0.88 | 0.56 | 0.83 | 0.95 | 0.89 | 0.75 | 0.81 | 0.89 | 0.82 | 0.74 |
| 0.84 | 0.59 | 0.8 | 0.95 | 0.9 | 0.72 | 0.81 | 0.92 | 0.82 | 0.71 |
| 0.83 | 0.61 | 0.91 | 0.94 | 0.92 | 0.79 | 0.76 | 0.89 | 0.81 | 0.68 |
| 0.84 | 0.58 | 0.88 | 0.97 | 0.86 | 0.74 | 0.78 | 0.87 | 0.88 | 0.73 |
| 0.85 | 0.63 | 0.75 | 0.96 | 0.88 | 0.75 | 0.65 | 0.88 | 0.84 | 0.71 |
| 0.83 | 0.64 | 0.81 | 0.92 | 0.94 | 0.8 | 0.73 | 0.92 | 0.9 | 0.79 |
| 0.85 | 0.58 | 0.81 | 0.94 | 0.87 | 0.72 | 0.7 | 0.89 | 0.9 | 0.63 |
| 0.85 | 0.55 | 0.86 | 0.95 | 0.91 | 0.69 | 0.74 | 0.9 | 0.86 | 0.72 |
| 0.86 | 0.58 | 0.83 | 0.94 | 0.89 | 0.79 | 0.8 | 0.86 | 0.83 | 0.71 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 40% | 0.84 | 0.54 | 0.79 | 0.93 | 0.84 | 0.68 | 0.72 | 0.87 | 0.81 | 0.7 |
| 0.82 | 0.56 | 0.83 | 0.93 | 0.83 | 0.74 | 0.7 | 0.86 | 0.77 | 0.7 |
| 0.85 | 0.56 | 0.85 | 0.97 | 0.89 | 0.78 | 0.72 | 0.86 | 0.79 | 0.7 |
| 0.83 | 0.54 | 0.86 | 0.96 | 0.93 | 0.77 | 0.77 | 0.93 | 0.81 | 0.7 |
| 0.83 | 0.56 | 0.77 | 0.94 | 0.83 | 0.74 | 0.74 | 0.86 | 0.83 | 0.64 |
| 0.85 | 0.6 | 0.82 | 0.93 | 0.88 | 0.72 | 0.71 | 0.86 | 0.82 | 0.71 |
| 0.83 | 0.55 | 0.8 | 0.93 | 0.91 | 0.77 | 0.73 | 0.86 | 0.84 | 0.74 |
| 0.84 | 0.58 | 0.8 | 0.95 | 0.86 | 0.7 | 0.72 | 0.83 | 0.86 | 0.72 |
| 0.82 | 0.57 | 0.85 | 0.95 | 0.91 | 0.79 | 0.78 | 0.89 | 0.86 | 0.66 |
| 0.88 | 0.53 | 0.83 | 0.99 | 0.89 | 0.75 | 0.71 | 0.91 | 0.8 | 0.72 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 30% | 0.76 | 0.56 | 0.86 | 0.91 | 0.87 | 0.73 | 0.68 | 0.91 | 0.81 | 0.74 |
| 0.82 | 0.56 | 0.78 | 0.92 | 0.88 | 0.64 | 0.76 | 0.83 | 0.81 | 0.72 |
| 0.81 | 0.58 | 0.86 | 0.95 | 0.84 | 0.76 | 0.76 | 0.85 | 0.9 | 0.67 |
| 0.86 | 0.59 | 0.81 | 0.96 | 0.89 | 0.72 | 0.68 | 0.84 | 0.83 | 0.73 |
| 0.83 | 0.59 | 0.81 | 0.89 | 0.9 | 0.68 | 0.71 | 0.87 | 0.81 | 0.72 |
| 0.69 | 0.54 | 0.81 | 0.91 | 0.9 | 0.7 | 0.77 | 0.88 | 0.81 | 0.7 |
| 0.8 | 0.58 | 0.76 | 0.91 | 0.88 | 0.68 | 0.66 | 0.89 | 0.77 | 0.7 |
| 0.83 | 0.48 | 0.82 | 0.87 | 0.85 | 0.72 | 0.66 | 0.84 | 0.78 | 0.69 |
| 0.83 | 0.58 | 0.73 | 0.9 | 0.84 | 0.7 | 0.69 | 0.88 | 0.81 | 0.7 |
| 0.81 | 0.54 | 0.82 | 0.92 | 0.82 | 0.68 | 0.68 | 0.86 | 0.77 | 0.66 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 20% | 0.84 | 0.59 | 0.73 | 0.86 | 0.82 | 0.66 | 0.73 | 0.85 | 0.88 | 0.65 |
| 0.8 | 0.57 | 0.76 | 0.85 | 0.83 | 0.74 | 0.68 | 0.74 | 0.84 | 0.62 |
| 0.75 | 0.47 | 0.79 | 0.89 | 0.77 | 0.68 | 0.66 | 0.82 | 0.76 | 0.66 |
| 0.79 | 0.51 | 0.75 | 0.91 | 0.81 | 0.66 | 0.61 | 0.87 | 0.78 | 0.65 |
| 0.71 | 0.51 | 0.73 | 0.89 | 0.84 | 0.66 | 0.67 | 0.77 | 0.77 | 0.64 |
| 0.76 | 0.52 | 0.74 | 0.82 | 0.84 | 0.65 | 0.61 | 0.84 | 0.73 | 0.61 |
| 0.74 | 0.51 | 0.72 | 0.87 | 0.8 | 0.72 | 0.62 | 0.83 | 0.82 | 0.66 |
| 0.81 | 0.44 | 0.82 | 0.85 | 0.84 | 0.68 | 0.66 | 0.81 | 0.78 | 0.62 |
| 0.84 | 0.45 | 0.73 | 0.91 | 0.87 | 0.63 | 0.65 | 0.83 | 0.72 | 0.68 |
| 0.72 | 0.51 | 0.78 | 0.91 | 0.82 | 0.64 | 0.62 | 0.81 | 0.73 | 0.6 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 10% | 0.79 | 0.5 | 0.65 | 0.84 | 0.75 | 0.58 | 0.63 | 0.68 | 0.69 | 0.51 |
| 0.71 | 0.43 | 0.61 | 0.77 | 0.78 | 0.67 | 0.59 | 0.71 | 0.77 | 0.57 |
| 0.71 | 0.49 | 0.64 | 0.83 | 0.79 | 0.62 | 0.57 | 0.74 | 0.65 | 0.6 |
| 0.74 | 0.42 | 0.66 | 0.82 | 0.78 | 0.66 | 0.54 | 0.82 | 0.63 | 0.56 |
| 0.76 | 0.45 | 0.68 | 0.87 | 0.75 | 0.7 | 0.58 | 0.72 | 0.76 | 0.57 |
| 0.73 | 0.45 | 0.68 | 0.88 | 0.83 | 0.58 | 0.57 | 0.76 | 0.71 | 0.58 |
| 0.7 | 0.47 | 0.7 | 0.8 | 0.8 | 0.64 | 0.56 | 0.75 | 0.69 | 0.58 |
| 0.67 | 0.39 | 0.61 | 0.83 | 0.76 | 0.58 | 0.57 | 0.71 | 0.65 | 0.54 |
| 0.7 | 0.44 | 0.64 | 0.77 | 0.71 | 0.57 | 0.58 | 0.69 | 0.72 | 0.5 |
| 0.76 | 0.33 | 0.7 | 0.81 | 0.72 | 0.6 | 0.58 | 0.67 | 0.73 | 0.56 |

HMM (DBtest)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 90% | 90 | 90 | 94 | 95 | 94 | 83 | 88 | 93 | 96 | 95 |
| 88 | 87 | 94 | 93 | 94 | 83 | 89 | 93 | 95 | 95 |
| 91 | 90 | 94 | 93 | 94 | 83 | 88 | 93 | 97 | 95 |
| 89 | 90 | 94 | 93 | 94 | 83 | 88 | 92 | 96 | 95 |
| 89 | 89 | 94 | 93 | 94 | 83 | 89 | 92 | 98 | 96 |
| 91 | 90 | 94 | 94 | 94 | 83 | 88 | 92 | 96 | 96 |
| 89 | 90 | 94 | 93 | 94 | 83 | 88 | 92 | 96 | 96 |
| 91 | 89 | 93 | 94 | 94 | 83 | 89 | 92 | 98 | 95 |
| 88 | 90 | 93 | 94 | 94 | 83 | 88 | 93 | 95 | 96 |
| 89 | 89 | 94 | 93 | 94 | 84 | 89 | 92 | 98 | 96 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 80% | 88 | 89 | 93 | 93 | 93 | 84 | 88 | 93 | 97 | 96 |
| 87 | 90 | 93 | 94 | 95 | 82 | 88 | 92 | 98 | 95 |
| 92 | 89 | 94 | 95 | 94 | 83 | 89 | 92 | 98 | 96 |
| 90 | 90 | 94 | 92 | 94 | 83 | 88 | 93 | 94 | 95 |
| 88 | 90 | 94 | 92 | 94 | 83 | 88 | 93 | 96 | 95 |
| 88 | 90 | 94 | 94 | 94 | 83 | 88 | 92 | 97 | 95 |
| 88 | 90 | 93 | 94 | 94 | 83 | 88 | 92 | 96 | 96 |
| 88 | 90 | 94 | 93 | 94 | 83 | 89 | 93 | 97 | 96 |
| 90 | 90 | 94 | 93 | 94 | 83 | 88 | 93 | 94 | 94 |
| 89 | 90 | 93 | 94 | 94 | 84 | 88 | 93 | 98 | 95 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 80% | 88 | 89 | 93 | 93 | 93 | 84 | 88 | 93 | 97 | 96 |
| 87 | 90 | 93 | 94 | 95 | 82 | 88 | 92 | 98 | 95 |
| 92 | 89 | 94 | 95 | 94 | 83 | 89 | 92 | 98 | 96 |
| 90 | 90 | 94 | 92 | 94 | 83 | 88 | 93 | 94 | 95 |
| 88 | 90 | 94 | 92 | 94 | 83 | 88 | 93 | 96 | 95 |
| 88 | 90 | 94 | 94 | 94 | 83 | 88 | 92 | 97 | 95 |
| 88 | 90 | 93 | 94 | 94 | 83 | 88 | 92 | 96 | 96 |
| 88 | 90 | 94 | 93 | 94 | 83 | 89 | 93 | 97 | 96 |
| 90 | 90 | 94 | 93 | 94 | 83 | 88 | 93 | 94 | 94 |
| 89 | 90 | 93 | 94 | 94 | 84 | 88 | 93 | 98 | 95 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 70% | 90 | 91 | 94 | 93 | 94 | 84 | 88 | 93 | 94 | 94 |
| 88 | 87 | 94 | 92 | 94 | 83 | 88 | 94 | 94 | 95 |
| 90 | 88 | 94 | 92 | 94 | 84 | 89 | 93 | 96 | 95 |
| 90 | 90 | 93 | 94 | 93 | 83 | 89 | 93 | 97 | 96 |
| 91 | 89 | 93 | 95 | 94 | 82 | 88 | 93 | 97 | 96 |
| 87 | 88 | 93 | 94 | 94 | 84 | 88 | 93 | 98 | 94 |
| 86 | 89 | 93 | 93 | 93 | 84 | 88 | 93 | 96 | 95 |
| 89 | 89 | 92 | 94 | 93 | 83 | 88 | 93 | 96 | 95 |
| 89 | 90 | 93 | 94 | 95 | 82 | 88 | 92 | 98 | 96 |
| 88 | 91 | 92 | 93 | 95 | 83 | 88 | 93 | 98 | 95 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 60% | 89 | 91 | 94 | 93 | 94 | 84 | 88 | 93 | 94 | 94 |
| 92 | 90 | 93 | 94 | 94 | 82 | 89 | 93 | 94 | 95 |
| 90 | 89 | 93 | 95 | 94 | 84 | 87 | 92 | 97 | 96 |
| 86 | 90 | 93 | 94 | 94 | 84 | 88 | 93 | 98 | 94 |
| 88 | 89 | 92 | 93 | 93 | 83 | 89 | 93 | 94 | 95 |
| 87 | 88 | 94 | 94 | 93 | 83 | 89 | 93 | 98 | 94 |
| 87 | 90 | 94 | 92 | 93 | 84 | 90 | 94 | 94 | 94 |
| 90 | 87 | 94 | 93 | 94 | 84 | 89 | 93 | 94 | 94 |
| 87 | 90 | 93 | 96 | 95 | 83 | 89 | 92 | 97 | 96 |
| 89 | 89 | 93 | 93 | 94 | 82 | 88 | 93 | 96 | 96 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 50% | 85 | 90 | 93 | 95 | 95 | 83 | 89 | 92 | 98 | 96 |
| 88 | 90 | 92 | 95 | 95 | 82 | 87 | 91 | 96 | 96 |
| 89 | 90 | 93 | 93 | 95 | 83 | 87 | 93 | 96 | 96 |
| 89 | 91 | 93 | 94 | 94 | 82 | 88 | 92 | 97 | 96 |
| 91 | 87 | 94 | 93 | 94 | 84 | 89 | 93 | 93 | 93 |
| 88 | 89 | 93 | 94 | 92 | 84 | 90 | 94 | 94 | 94 |
| 88 | 89 | 94 | 92 | 93 | 84 | 89 | 96 | 94 | 94 |
| 89 | 88 | 94 | 96 | 93 | 84 | 89 | 93 | 96 | 94 |
| 88 | 88 | 93 | 96 | 94 | 84 | 89 | 93 | 98 | 94 |
| 89 | 89 | 94 | 94 | 93 | 82 | 89 | 91 | 94 | 94 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 40% | 89 | 88 | 94 | 92 | 94 | 82 | 89 | 91 | 92 | 93 |
| 87 | 88 | 94 | 93 | 93 | 83 | 91 | 95 | 94 | 95 |
| 88 | 90 | 93 | 95 | 94 | 85 | 89 | 93 | 96 | 94 |
| 89 | 90 | 93 | 94 | 94 | 84 | 88 | 92 | 98 | 95 |
| 91 | 89 | 92 | 94 | 94 | 83 | 87 | 81 | 96 | 95 |
| 90 | 88 | 94 | 93 | 94 | 83 | 88 | 92 | 96 | 96 |
| 88 | 88 | 94 | 93 | 93 | 85 | 90 | 93 | 94 | 92 |
| 90 | 88 | 94 | 95 | 93 | 84 | 89 | 92 | 96 | 94 |
| 90 | 88 | 94 | 93 | 94 | 83 | 88 | 92 | 96 | 96 |
| 87 | 88 | 94 | 94 | 95 | 85 | 88 | 93 | 97 | 95 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 30% | 89 | 88 | 93 | 93 | 93 | 83 | 87 | 90 | 93 | 92 |
| 90 | 89 | 94 | 92 | 92 | 84 | 89 | 91 | 93 | 94 |
| 88 | 90 | 93 | 93 | 92 | 85 | 91 | 94 | 94 | 93 |
| 88 | 88 | 94 | 94 | 93 | 84 | 90 | 92 | 96 | 93 |
| 89 | 89 | 94 | 94 | 93 | 84 | 89 | 93 | 95 | 93 |
| 88 | 90 | 93 | 95 | 94 | 85 | 89 | 94 | 96 | 95 |
| 90 | 90 | 93 | 96 | 94 | 82 | 87 | 94 | 97 | 95 |
| 87 | 91 | 92 | 94 | 96 | 83 | 88 | 91 | 96 | 96 |
| 88 | 89 | 90 | 93 | 96 | 83 | 86 | 93 | 92 | 94 |
| 91 | 86 | 91 | 93 | 94 | 83 | 86 | 91 | 93 | 95 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 20% | 91 | 88 | 93 | 94 | 94 | 84 | 85 | 88 | 91 | 94 |
| 92 | 86 | 92 | 93 | 93 | 83 | 87 | 92 | 95 | 93 |
| 88 | 89 | 91 | 93 | 92 | 81 | 90 | 92 | 94 | 93 |
| 88 | 89 | 95 | 94 | 94 | 85 | 89 | 92 | 96 | 92 |
| 89 | 89 | 94 | 93 | 95 | 83 | 90 | 96 | 96 | 93 |
| 88 | 90 | 94 | 93 | 93 | 82 | 86 | 94 | 94 | 95 |
| 87 | 91 | 92 | 96 | 94 | 83 | 86 | 93 | 94 | 94 |
| 86 | 89 | 93 | 95 | 94 | 81 | 84 | 93 | 93 | 94 |
| 89 | 90 | 92 | 94 | 96 | 82 | 82 | 91 | 94 | 96 |
| 88 | 85 | 92 | 95 | 95 | 83 | 84 | 91 | 91 | 92 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speaker  Whi\_Spe subset | 2 | 6 | 7 | 8 | 9 | 14 | 15 | 16 | 18 | 19 |
| 10% | 89 | 86 | 92 | 94 | 93 | 81 | 85 | 87 | 88 | 91 |
| 87 | 89 | 88 | 92 | 93 | 77 | 82 | 87 | 90 | 90 |
| 88 | 89 | 90 | 93 | 92 | 80 | 83 | 92 | 89 | 89 |
| 86 | 89 | 90 | 90 | 92 | 81 | 89 | 90 | 95 | 92 |
| 89 | 86 | 93 | 94 | 93 | 77 | 88 | 92 | 90 | 92 |
| 84 | 89 | 92 | 90 | 81 | 83 | 85 | 92 | 97 | 88 |
| 85 | 87 | 94 | 94 | 93 | 82 | 84 | 91 | 90 | 90 |
| 86 | 87 | 90 | 94 | 92 | 80 | 84 | 91 | 95 | 91 |
| 87 | 88 | 85 | 96 | 94 | 75 | 83 | 89 | 84 | 93 |
| 81 | 86 | 89 | 92 | 94 | 80 | 81 | 88 | 91 | 91 |

100% HMM

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 101 | 102 | 103 | 104 | 105 | 109 | 111 | 112 | 113 | 115 |
| 90 | 88 | 94 | 93 | 94 | 83 | 88 | 92 | 96 | 96 |

**Table 2.** Average recognition accuracy (DBtest database) for 7 data augmentation techniques and HMM framework.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Augmentation | | | | | | | |
| Training  scenario | PS | TST | VC | PS+TST | PS+VC | TST+VC | PS+TST+VC |
| W | 90.66 | 90.66 | 91.61 | 89.06 | 90.75 | 90.98 | 89.88 |
| P | 72.16 | 71.32 | 69.51 | 72.11 | 72.06 | 72.26 | 70.84 |
| OW | 91.29 | 91.21 | 91.61 | 91.09 | 91.29 | 91.49 | 91.14 |
| OP | 91.17 | 90.91 | 91.34 | 91.06 | 91.20 | 91.09 | 90.88 |
| OWP | 91.53 | 91.91 | 92.09 | 91.37 | 91.48 | 92.06 | 91.52 |
| ORIGINAL  (NO AUG) | 91.61 | | | | | | |

**Table 3.** Average recognition accuracy (DBtest database) for 7 data augmentation techniques and CNN framework.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Augmentation | | | | | | | |
| Training  scenario | PS | TST | VC | PS+TST | PS+VC | TST+VC | PS+TST+VC |
| W | 84.29 | 83.72 | 83.88 | 84.22 | 84.37 | 83.70 | 84.93 |
| P | 69.52 | 69.76 | 69.23 | 69.94 | 69.00 | 69.72 | 68.85 |
| OW | 85.82 | 84.98 | 85.18 | 85.62 | 85.84 | 85.02 | 85.80 |
| OP | 85.73 | 85.80 | 86.17 | 86.10 | 85.74 | 85.16 | 85.80 |
| OWP | 87.54 | 87.37 | 87.28 | 87.97 | 88.06 | 86.86 | 87.31 |
| ORIGINAL  (NO AUG) | 84.14 | | | | | | |