Rangkuman

Building a Speech and Text Corpus of Turkish : Large Corpus Collection with Initial Speech Recognition Results

Fungsi utama dari Automatic Speech Recognition (ASR) yaitu mengubah pengucapan manusia menjadi teks tertulis. Implementasi ASR dapat digunakan pada data social media, kalimat perintah untuk perangkat, Subtitle dari film, dll. ASR bekerja dengan model yang dihasilkan dari pelatihan machine learning berdasarkan klasifikasi pola statistik. ASR menggunakan supervised learning untuk melatih speech classifier. Pelatihan ASR ini menggunakan metode Gaussian Mixture Model – Hidden Markov Models (GMM-HMM). Meskipun metode ini sudah sukses diterapkan tetapi akurasi ASR masih terbatas untuk kemampuan manusia berbicara.

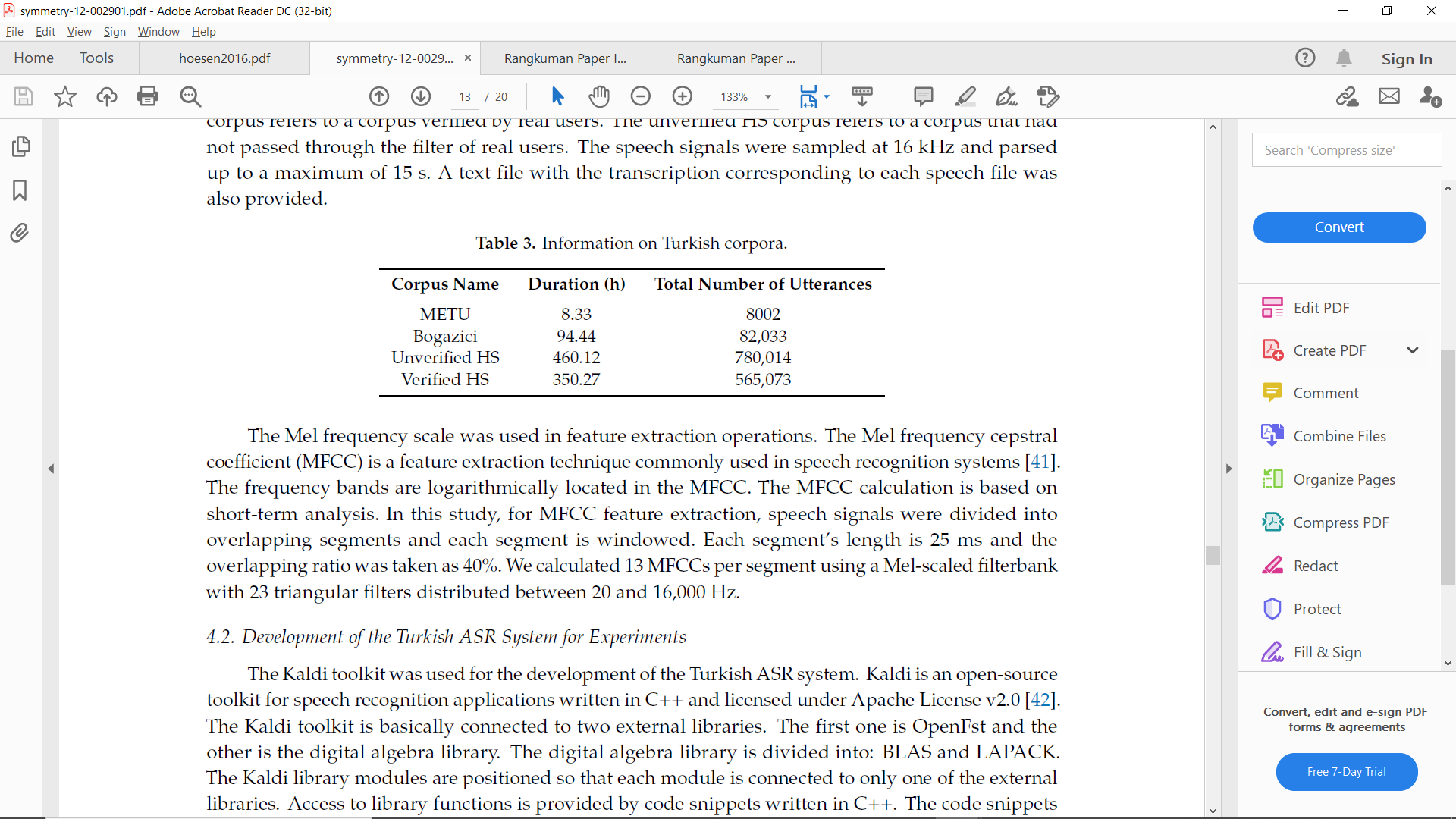
Sebelum melakukan implementasi ASR perlu dilakukan pengumpulan data speech corpus. Metode yang digunakan untuk mengumpulkan data speech corpus sebagai berikut :

1. Use of Film Movies and Time-Bound Subtitle Documents (subtitle film)
2. Collections of Speech Data with a Mobile Application (google translate text to speech)
3. Collections of Speech Data with Transfer Learning (data berasal dari machine learning lainnya)

Langkah-langkah implementasi ASR

1. Speech Corpus for Experiments

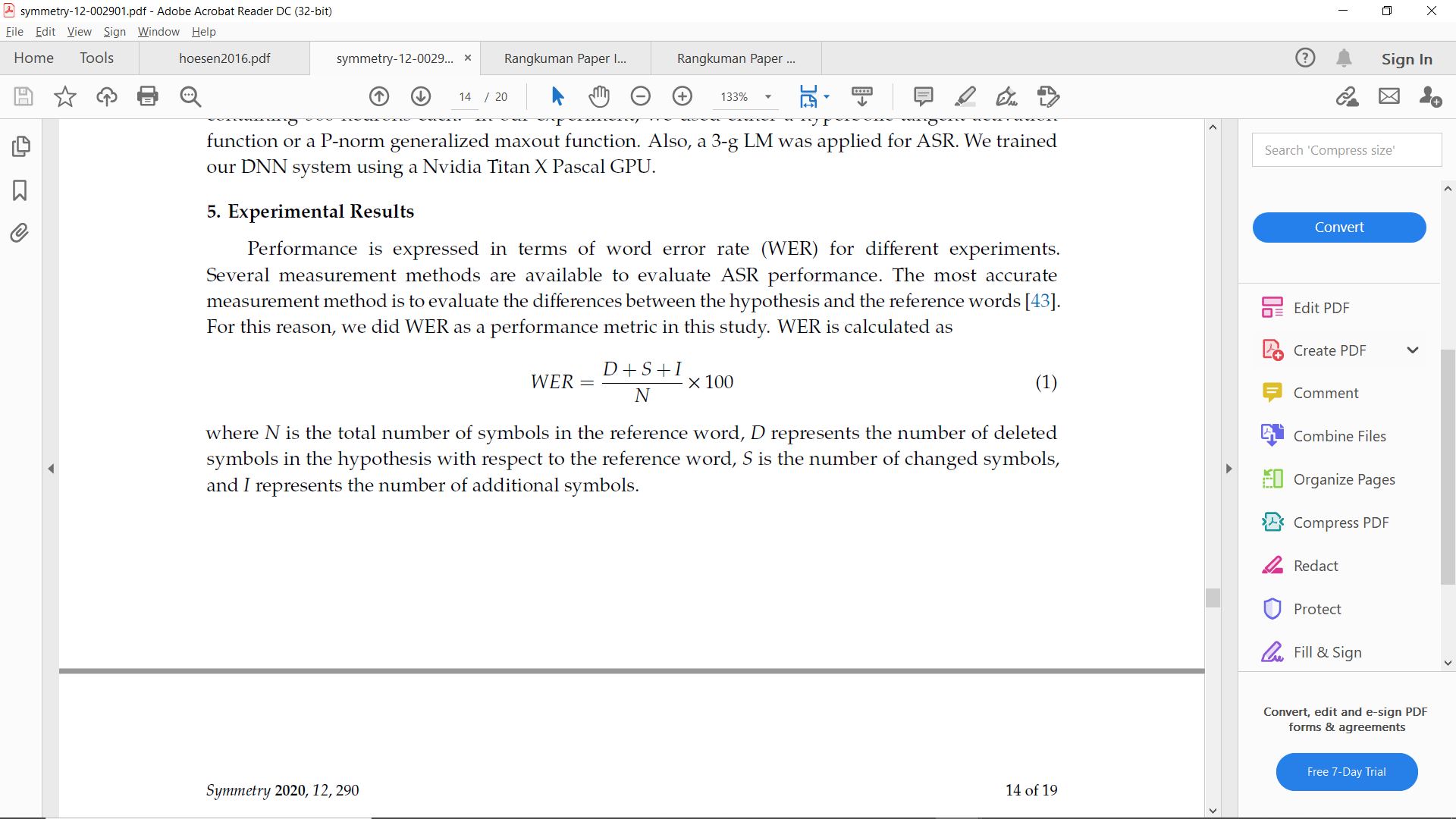
Corpus mengandung banyak variasi suara dengan frekuensi sampling 16 kHz dan 16 bit sample size



1. Development of the Turkish ASR system for experiments

Toolkit Kaldi digunakan untuk pengembangan sistem ASR Turki. Kaldi merupakan toolkit open source digunakan untuk aplikasi ASR dan ditulis dalam Bahasa pemrograman C++. Model ASR menggunakan GMM-HMM dan Deep Neural Network (DNN).

Performa dari ASR ini dapat diukur menggunakan Word Error Rate (WER).

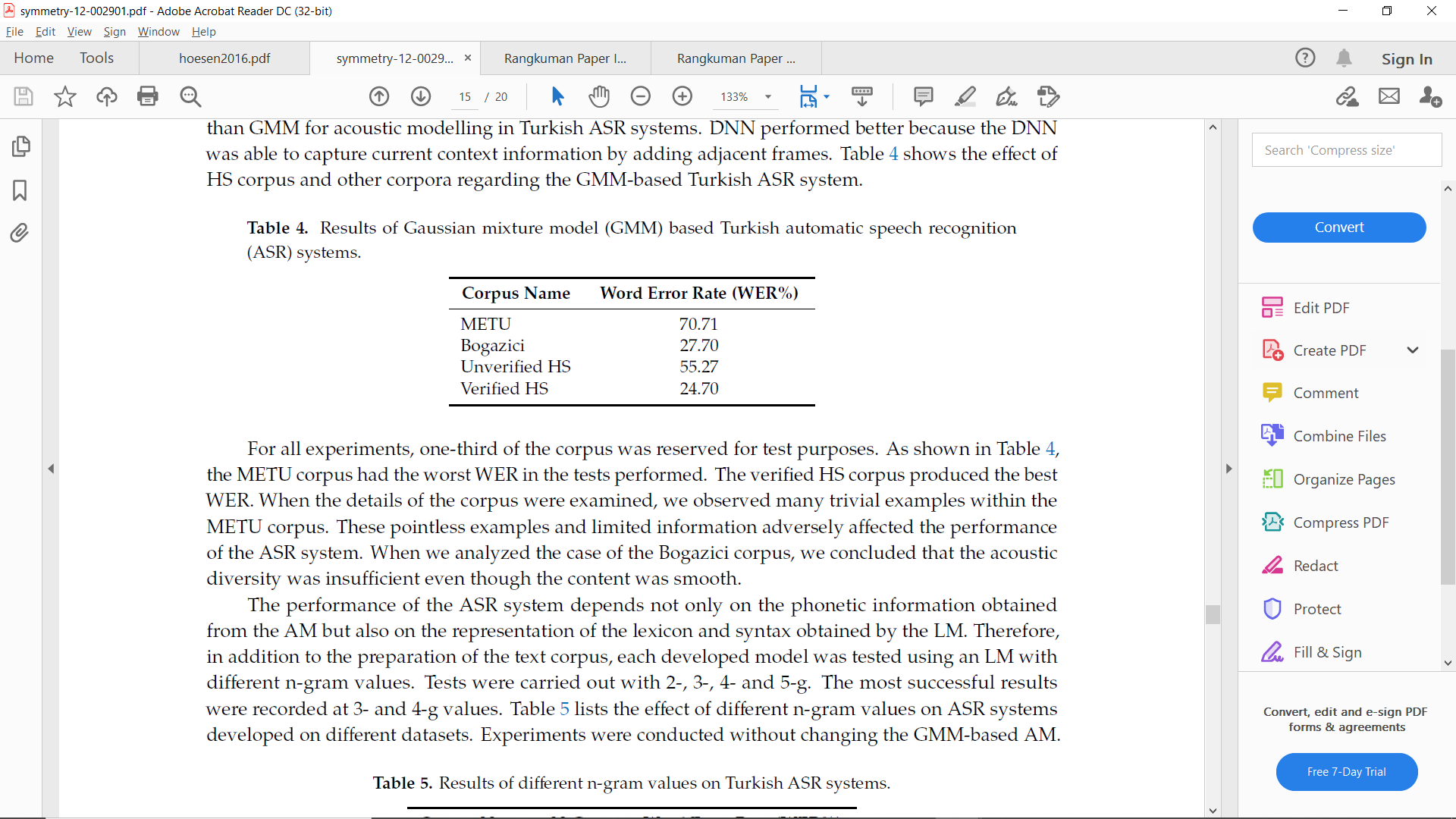


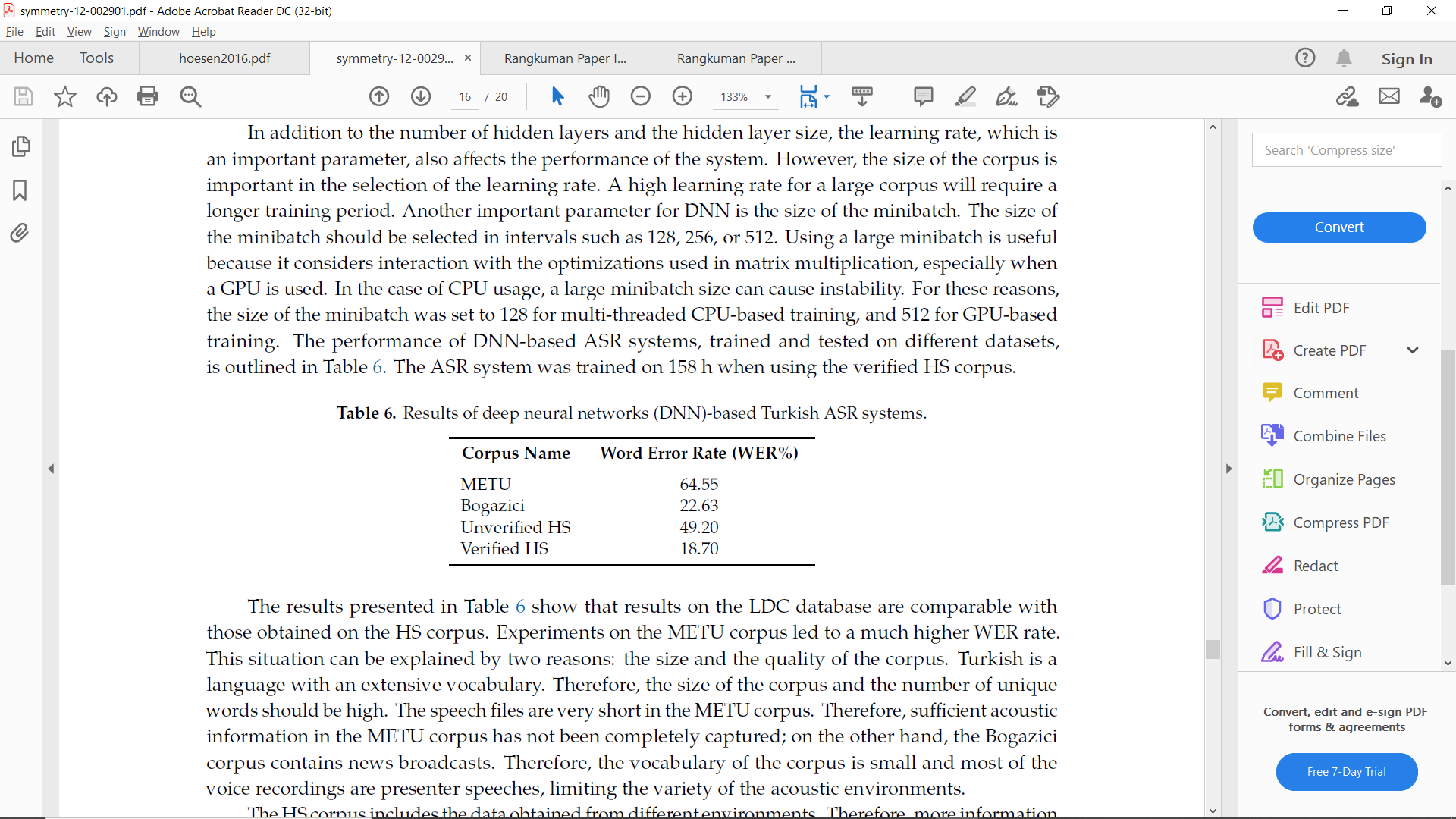
S = Substitution

I = Insertion

D = Deletion

N = Jumlah Kata pada test set

Dari hasil experiment didapatkan data sebagai berikut



Dari kedua data diatas dapat disimpulkan bahwa model akustik DNN memberikan performa yang lebih baik dibandingkan GMM-HMM karena WER yang semakin kecil.