

Appendix for Sensors to Sign Language: A Natural Approach to Equitable Communication

Submission 1661

February 2021

1 Dataset vocabulary

Our dataset consists of 558 distinct words, which were selected based on being the most popular American Sign Language words. Our lexicon, in alphabetical order, is as follows:

abandon, abbreviate, about, above, accept, access, accident, accidental, accounting, across, active, activity, add, addictive, address, admire, admit, adventure, advertise2, advice2, advocacy, affiliate, afraid, africa, after, afternoon, again, against, age, aggressive, agree, ahead, airplane, alarm, all, all-right, allergy, allow, almost, alone, also, altogether, always, amazing, ambulance, america, analyze, and, angry, animal, animation, announce, annual, another, answer, any, apart, apology, appearance, apple, appoint, approach, approve, area, argue, arm, armadillo, army, arrest, arrive, art, article, as-since, ask, assist, assistant, association, asthma, attention, attitude, attract, auction, audiology, aunt, australia, australia2, austria, austria2, automatic, average, avoid, awake, award-verb, baby, bacon, bad, bag, barely, basic, basketball, bath, battery, beaver, bed, beer, believe, best, better, bicycle, bird, birthday, black, body, book, bored, borrow, boss, both, bother, bowl, box, boy, boyfriend, bread, break, bright, bring, brother, bus, business, busy, but, butter, buy, calendar, california, calm, camera, can, cannot, caption, careful, casual, cat, category, celebrate, cereal, chair, champion, chaos, charge, chase, cheap, check, cheese, cherish, child, children, chocolate, clash, clean, clothes, coffee, college, color, come, comfortable, communication, community, complain, confident, confidential, conflict, contact, convince, cook, cookie, cost, crewcut, criticize, culture, cup, curious, daughter, delicious, different, dirty, disappoint, disconnect, discuss, dislike, document, dominant, door, double, doubt, drink, drive, dry, each, egg, eleven, email, emotion, end, energy, enjoy, enough, enter, environment, escape, evaluate, evening, event, everything, example, excited, expect, expensive, experience, eyes, face, family, father, fault, feel, find, fingerspell, flower, for, forest, former, friday, friend, from, fruit, full, fun, funny, future, game, generation, get, go, good, grass, green, group, grow, guide, gum, habit, half, hands, happen, happy, hard, hate, have, he, headache, hearing, hello, help, here, hide,

history, home, homework, honest, hour, house, how, hug, hungry, hurry, hurt, husband, idea, impact, important, impress, improve, in, independent, inside, institute, insurance, interested, internet, interpreter, iphone, join, joke, journey, keep, kiss, kitchen, know, language, laptop, late, later, laugh, law, lazy, learn, library, lie, like, list, lock, long, lousy, love, mad, mail, make, man, many, maybe, me, medicine, meet, meeting, member, mess, middle, milk, minute, mirror, monday, money, month, more, morning, most, mother, move, movie, much, must, my, myself, name, near, need, never, new, next, nice, night, no, nonsense, normal, north, not, nothing, now, obvious, off, office, often, ok, on, once, onion, open, opportunity, option, orange, other, our, outside, over, overlook, overnight, overwhelmed, paper, pay, pen, people, permission, person, perspective, picture, place, plan, plant, plate, play, popular, positive, possible, postpone, potato, practice, precious, predict, pretty, print, priority, program, promise, propose, provide, purple, put, quiet, quit, rain, read, really, reason, red, reduce, refuse, register, regular, relationship, relax, remember, remind, resist, rest, restaurant, road, room, ruin, rule, sad, safe, salt, saturday, schedule, school, science, see, seem, select, sensitive, serious, share, shoes, should, show, sick, signature, silly, since, sister, sit, situation, skill, sleep, small, smart, smell, smooth, social, some, something, sometimes, sorry, space, star, starbucks, start, stay, still, strange, street, stress, struggle, stuck, study, summer, summon, sure, sweet, table, take, talk, taste, teacher, team, tease, technology, telephone, tell, temporary, test, text, think, thirsty, thrilled, thursday, tired, today, toilet, tomato, tomorrow, toothbrush, topic, touch, tough, trash, tree, trouble, truth, tuesday, understand, universe, university, use, vacation, vegetable, village, vision, visit, vlog, volunteer, wait, want, warn, watch, we, wednesday, week, weight, when, which, who, why, wife, win, window, wine, winter, wise, wish, with, woman, wonder, word, work, world, wow, write, yay, yeah, year, yes, yesterday, you, young, your, yourself

2 Common Misclassifications

The 20 most commonly misclassified words are:

boy, chase, study, delicious, internet, tell, toilet, convince, strange, opportunity, room, birthday, precious, member, late, confident, nice, window, some, world

3 Hyperparameters

Our code, which we will make publicly available, runs typical deep learning architectures: LSTM, BiLSTM, CNN, and a BiLSTM+CNN.

As mentioned in the paper, we varied the parameters optimized our hyperparameters on the development set by uniformly sampling, varying the number of hidden layers from 1 to 3, the number of hidden units from 64 to 1024, the

number of epochs from 10 to 300, dropout from 0 (no dropout) to 0.4. We used Adam as our optimizer.

We did not exhaustively perform this search, as it became clear that most values gave comparable, strong results without too much variance. In total, we ran roughly 50-100 individual experiments during a 2 day span.

Our final system that yielded the best results for us were:

- 1 BiLSTM layer with 768 units
- 1 BatchNorm layer
- 1 dropout layer w/ 0.4
- 1 1DConv w/ 64 kernels and kernel size of 4 (ReLU activation)
- 1 BatchNorm layer
- 1 MaxPooling1D layer
- 1 dropout layer w/ 0.4
- 1 1DConv w/ 16 kernels and kernel size of 2 (ReLU activation)
- 1 BatchNorm layer
- 1 MaxPooling1D layer
- 1 Dense layer of size 558

This resulted in 5,451,038 parameters. Adam was our optimizer. We ran our model for 250 epochs. It ran in 30 minutes on a free Google Colab (GPU-enabled) notebook. When using an Ensemble approach, we altered which of the 5 non-test signs would serve as the validation set (a la cross-validation). However, at each time, we produced predictions on the test set. Our assembled approach averaged all of these predictions in order to make a final prediction. Thus, this model was a composition of 5 individual BiLSTM+CNN runs, so it took 5 times as long to run as a single BiLSTM+CNN model does. It increased accuracy by roughly 1%.