



Universidad Nacional Autónoma de México

Facultad de ingeniería

Materia: Modelos de Programación Orientada a Objetos

Grupo: 04

Proyecto: “Hablando Juntos”

Profesor: Germán Santos Jaimes

Integrantes:

Hernández Urbina José Abraham

Luna Roa Miguel Ángel

Muñoz Ponce Daniel Sebastián

Semestre 2019-2

Propósito y descripción

El presente documento describe las metas, el proceso de creación y logros alcanzados por nuestra aplicación aún en desarrollo; “Hablando Juntos”.

Hablando Juntos es una aplicación la cual busca una forma eficiente y sencilla de comunicarse con personas que tengan alguna discapacidad del habla, esto mediante el aprendizaje del lenguaje de señas, gracias al *Machine Learning* (CoreML) es que desarrollamos un modelo el cual logra distinguir entre las diferentes señas propias del lenguaje.

1.0 Proceso creativo, nacimiento de la idea y modelado de la misma.

Buscábamos desarrollar una aplicación la cual fuera “ruidosa”, que hablara por sí misma y fuera una herramienta para la unión, por lo que pensamos: “¿Qué mejor que la educación?”. Así es como surgieron distintas ideas, herramientas para el estudio, administradores de trabajos, entre otros.

Hablando Juntos paso por una serie de transformaciones desde su concepción, desde el nombre hasta el concepto mismo. Primeramente la aplicación se había pensado como un traductor de braille a español, a esta aplicación la llamábamos “Sintiendo Juntos”. Sin embargo, las limitaciones del hardware nos condujeron a reinventar la aplicación, explorar otras alternativas para que la comunicaciones fuera eficiente entre todas las personas sin importar las diferencias dialécticas que las pudieran separar, específicamente personas que, por alguna razón carecen de algún sentido el cual les impide la comunicación. Aquí es donde Hablando Juntos nació, nuestra aplicación ahora se centraría en la traducción del lenguaje de señas al español.



Nuestra idea inicial consistía en crear un traductor, el cual con apoyo de la cámara estaría en la posición de traducir todas las expresiones emitidas por una persona, posteriormente en una caja de texto el receptor tendría a su disposición la traducción de lo antes grabado. Surgieron en ese momento surgieron las preguntas “¿Cuál sería la mejor forma de hacer un traductor de tal índole?”. La respuesta era sencilla, pero difícil de implementar o eso parecía; agregar Machine Learning a nuestra aplicación, para así reconocer las señas y dar una traducción.

1.1 Aprendiendo sobre Machine Learning y descubriendo CoreML

Una vez resuelto el dilema en torno a la estructura base de la aplicación procedimos con la investigación del Machine Learning, por desgracia existe poca información sobre el tema, y existen un número menor de proyectos basados en esta estructura .

Sin embargo, con ayuda del profesor encontramos una “reciente” propuesta de Apple la cual denominan CoreML, un par de líneas de código las cuales permiten crear un simple entrenador de modelos estáticos. Aun así necesitábamos explorar más el campo del Machine Learning, por lo cual seguimos investigando y encontramos el repositorio: <https://github.com/likedan/Awesome-CoreML-Models> el cual contiene muy buena documentación acerca de CoreML el cual nos ayudó a comprender la lógica detrás de los modelos así como su manipulación.

Encontramos artículos y sitios de los cuales nos apoyamos para realizar nuestra aplicación, uno de ellos es el sitio oficial de Apple Developer el cual pone a la disposición del desarrollador documentación concisa y eficiente, así como el portal AppCoda el cual nos facilitó la comprensión lógica detrás del manejo de los modelos ya entrenados.



2.0 Creación de la interfaz de usuario

El desarrollo de la interfaz de usuario fue sencilla gracias a las herramientas propias de Xcode, este entorno nos permitió una gran variedad y libertad a la hora de realizar la parte estética de la aplicación. Desde el comienzo establecimos como prioridad tener un modelo eficiente y con un bajo margen de error, por esto nos enfocamos en un ochenta por ciento en el desarrollo del modelo y optamos por crear una estética simple pero funcional para la aplicación.

El primer diseño de nuestra aplicación solamente consideraba los botones y vistas destinados a la traducción, sin embargo, optamos por proponer más funciones las cuales aumentarían el atractivo de nuestro proyecto. Es aquí donde integramos la parte del aprendizaje, los apartados principales contemplaban: una sección para que el usuario visualizará el alfabeto en el lenguaje de señas, una sección sobre los diferentes artículos del hogar, una sección que le

permitiría a las persona con esta discapacidad poder reproducir audios para expresar sus necesidades.



Vocales



Abcdario



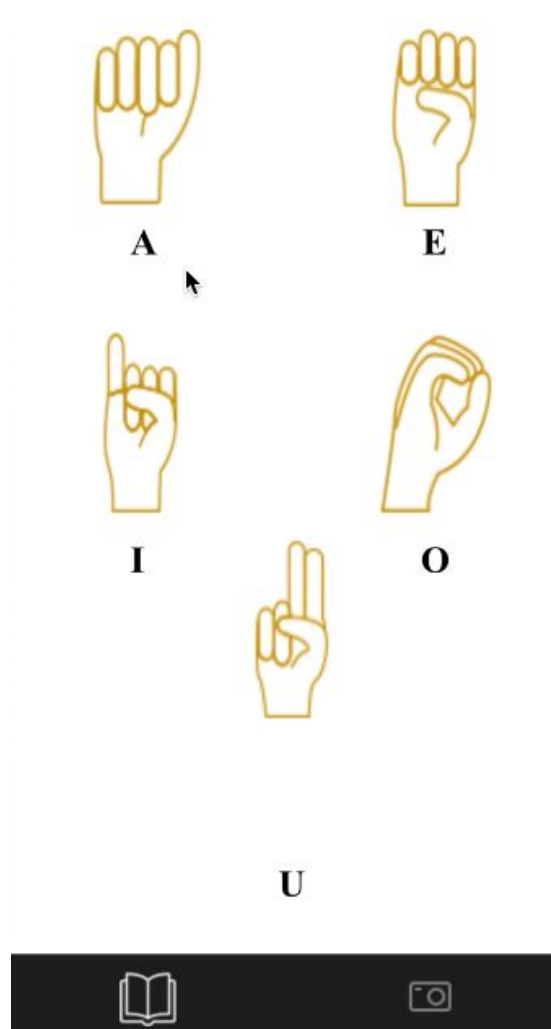
Cosas del Hogar



Galería

2.1 Opción de aprendizaje

La opción de aprendizaje contaría con ventanas destinadas a introducirte al lenguaje de señas, seguido de práctica impulsada por CoreML, lo cual permitiría un desarrollo integral del usuario al aprender y practicar simultáneamente. Está vista a su vez contaría con opciones como “abecedario” y “vocabulario”, las cuales darían la traducción en lenguaje de señas y en español.



2.2 Sección cosas del hogar

Esta sección serviría de ejemplo para un futuro seccionar toda la información y hacer más amable y sencillo el aprendizaje, tomando todas las palabras y oraciones que pudieran tener relación con el lugar o ámbito relacionado, en este caso el hogar. Sin embargo, este apartado se mantuvo como una idea para desarrollar a futuro debido al escaso tiempo del cual disponíamos para el desarrollo del proyecto.



Cosas del Hogar

3.0 Nuestros límites

Los límites en el proyecto fueron claros y podemos resumirlos en dos factores: capacidad de procesamiento y tiempo.

3.1 Capacidad de procesamiento, una desventaja clara

Estábamos entusiasmados por la creación de un modelo poderoso, capaz de diferenciar entre las distintas configuraciones de las manos en el lenguaje de señas, además de darnos una traducción coherente, sin embargo encontramos nuestro primer límite en el desempeño de las computadoras, las cuales se rehusaban a cumplir nuestra demandante labor y protestaban con claras señales; el sobre calentamiento de las unidades, un requerimiento excesivo de recursos el cual se veía reflejado en la pérdida progresiva de batería, así como también un ralentizamiento bastante notorio de las demás tareas mientras se realizaba el entrenamiento del modelo.

3.2 Tiempo

El tiempo destinado a la aplicación fue justo, sin embargo, la gran cantidad de datos a reunir era amplia y no contábamos con el tiempo necesario para crear un traductor completo, esto fue otro punto por lo cual decidimos replantear la meta y reducirla a la enseñanza de las vocales.

3.3 Aplicación limitada

Al investigar más acerca del funcionamiento de los modelos de reconocimiento de imágenes basados en Machine Learning nos dimos cuenta de que sería muy complicado alcanzar todas las metas planteadas en un principio, discutimos en equipo sobre los aspectos más importantes, decidimos enfocarnos en tener un modelo bien entrenado el cual nos permitiera alcanzar una lectura con una precisión mayor al ochenta por ciento.

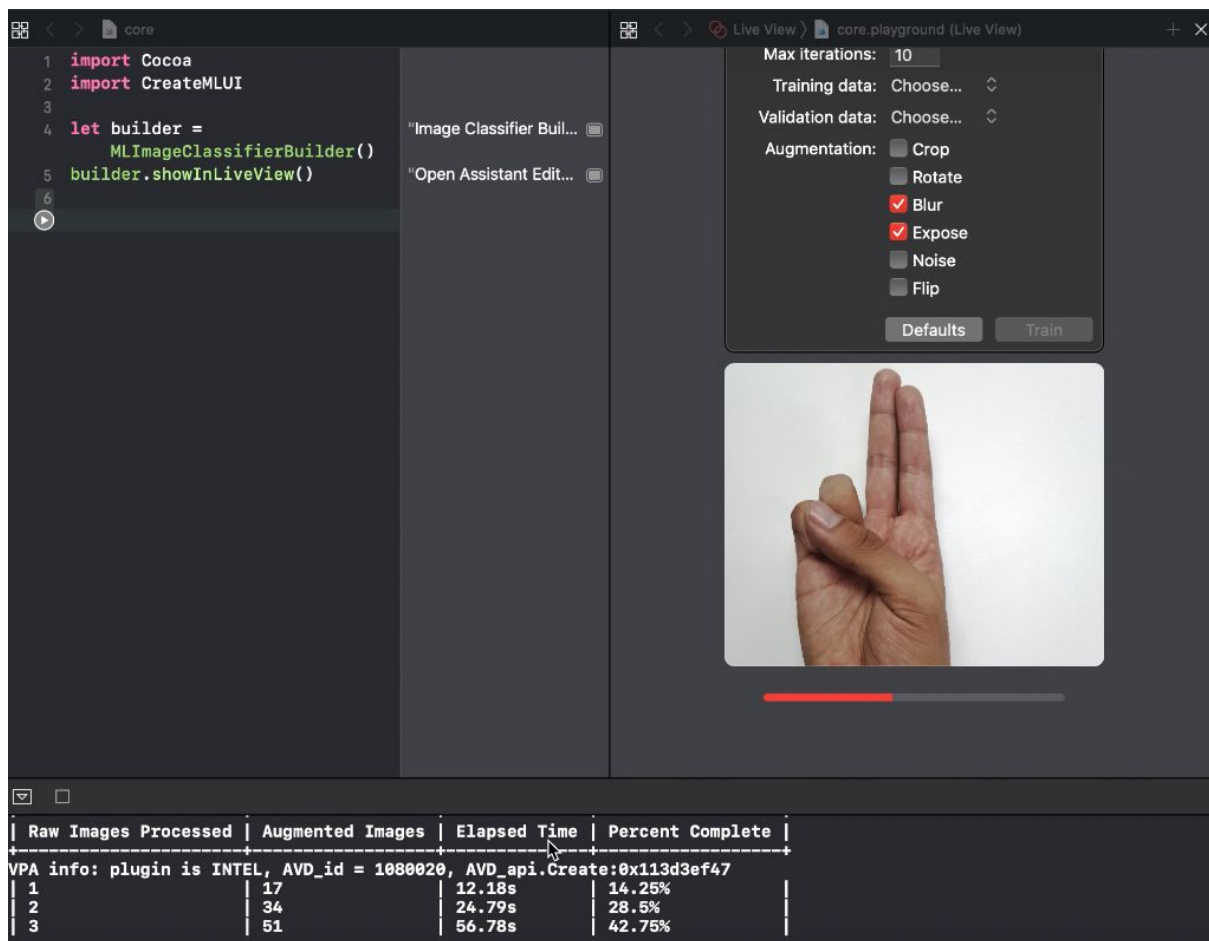
Nuestra aplicación ahora se encontraba reducida, pasamos de la idea de un traductor completo a un traductor sencillo de las vocales. Nuestras metas a corto plazo había cambiado, nuestro nuevo panorama presentaba la oportunidad de crear un modelo más sólido. Durante las últimas semanas nos dedicamos a mejorar, reunir y filtrar las fotografías las cuales servirían para el modelado.

Una nueva meta fue marcada; tener una aplicación la cual mostrara al usuario la configuración de las manos para cada una de las vocales en lenguaje de señas. La aplicación pondría a disposición del usuario la configuración de las manos para cada una de las vocales, posteriormente podría comprobar lo aprendido tomando una fotografía y es ahí donde la eficiencia de nuestro modelo nos permitiría interpretar correctamente la seña si es que esta fue bien ejecutada y capturada.

4.0 Creando el Modelo

Aunque programáticamente la creación del modelo resulta sencilla, el modelo en sí no lo es. El modelado implica la recolección de datos precisos la cual se vuelve increíblemente extensa, es necesario cubrir todas o la mayoría de variantes que puedan presentarse para una determinada situación. En nuestro caso fue recolectar las fotografías de las diferentes posiciones de las manos para cada una de las vocales.

Se recolectaron 350 fotografías de diferentes personas, nos enfocamos en lograr un entorno similar para todas las fotografías esto con la finalidad de lograr un mejor modelado lo cual resultaría en una mejor lectura, el modelo fue entrenado dentro de los estándares que nos permitieron las unidades del laboratorio, esperábamos desarrollar un entrenamiento más riguroso y aunque esto no fue posible logramos desarrollar un modelo el cual se encontraba dentro de nuestros estándares.



4.1 Versión Beta

Hablando juntos es capaz de diferenciar de entre cinco diferentes configuraciones de manos hasta ahora, con una precisión de reconocimiento mayor al setenta y cinco por ciento, cuenta con una sección que le permite al usuario visualizar las cinco diferentes posiciones de manos. Esperamos contar con el apoyo del laboratorio para seguir desarrollando nuestra aplicación, tenemos planeado a corto plazo agregar las opciones de *diccionario* y *cosas del hogar*.

Al contar con las opciones que se encuentran en desarrollo habilitadas nos enfocaremos en cumplir con las especificaciones marcadas por Apple para colocar nuestra aplicación junto a las miles de aplicaciones existentes en la App Store y así, con ayuda de los usuarios lograremos incrementar la base de datos para seguir entrenando nuestro modelo y lograr aumentar su eficiencia.

Conclusiones.

Hablando juntos es un proyecto amplio y ambicioso el cual busca educar, unir y concientizar sobre las discapacidades que nos impiden comunicarnos libremente, nos llama a ser igualitarios con aquellos quienes fueron privados de algún sentido. Hablando juntos significó un reto, no sólo en el aspecto de la programación, si no de conciencia y conocimiento general, nos indujo a empaparnos de una cultura que vaga junto a la nosotra pero es poco valorada; aprender lenguaje de señas. Tuvimos que realizar una investigación sobre el lenguaje de señas en México ya que existen diferencias en el lenguaje dependiendo del país e idioma.

Hablando juntos es una aplicación que necesita tiempo y dedicación, necesita ser pulida para así alcanzar los estándares del AppStore, además, necesita del tiempo y dedicación para cumplir con las metas que fueron marcadas por el equipo.