

UNIVERSIDAD NACIONAL DE  
SAN CRISTÓBAL DE HUAMANGA

FACULTAD DE INGENIERÍA DE MINAS, GEOLOGÍA Y CIVIL

UNIDAD DE INVESTIGACIÓN E INNOVACIÓN  
PROGRAMA: INGENIERÍA DE SISTEMAS  
ÁREA: CIENCIA COMPUTACIONAL Y MÉTODOS NUMÉRICOS



PROYECTO DE INVESTIGACIÓN

**Método para Reconocimiento de Gestos Aislados de la Lengua de Señas en Tiempo Real, Área de Salud, 2019**

**RESPONSABLE** : Mg. Ing. Manuel Avelino Lagos Barzola

**MIEMBROS** : MSc. Ing. Efraín Elías Porras Flores  
MSc. Ing. José Elías Yauri Vidalón

**COLABORADOR** : Ing. Lisber Arana Hinojosa

AYACUCHO – PERÚ

2019

**Fecha de Inicio: 01 de enero 2019 - Fecha de Término: 31 de diciembre 2019**

## ÍNDICE

	Pág.
I. GENERALIDADES	3
1.1. Título	3
1.2. Responsable	3
1.3. Resumen	3
1.4. Tipo de Investigación	3
1.5. Cronograma	3
1.6. Recursos disponibles	4
1.7. Presupuesto	4
1.8. Financiamiento	5
II. PLAN DE INVESTIGACIÓN	5
2.1. Problema	5
2.2. Objetivos	5
2.3. Marco teórico	6
2.3.1. Antecedentes de la investigación	6
2.3.2. Gestos Aislados de la Lengua de Señas	6
2.3.3. Tiempo Real	7
2.4. Hipótesis	7
2.5. Variables e indicadores	7
2.6. Diseño metodológico	8
2.7. Referencias bibliográficas	8
III. ANEXOS	10
3.1. Matriz de consistencia	10

## **I. GENERALIDADES**

### **1.1. TITULO**

Método para Reconocimiento de Gestos Aislados de la Lengua de Señas en Tiempo Real, Área de Salud, 2019.

**1.2. RESPONSABLE:** Mg. Ing. Manuel Avelino Lagos Barzola.

#### **MIEMBROS:**

MSc. Ing. Efraín Elías Porras Flores

MSc. Ing. José Elías Yauri Vidalón

**COLABORADOR:** Ing. Lisber Arana Hinostroza

### **1.3 RESUMEN**

Mientras la mayoría de seres humanos se comunican utilizando alguna lengua oral (LO), las personas sordo mudas utilizan la lengua de señas (LS) como sistema de comunicación. A diferencia de las LOs que utilizan sonidos articulados, las LSs utilizan movimientos estandarizados de las manos, cuerpo, junto con expresiones faciales para transmitir mensajes. Cualquier idea, concepto, emoción, sentimiento, pueden ser comunicado mediante lengua de señas (Cooper, Holt, & Bowden, 2011).

Sin embargo, las personas sordas enfrentan serios problemas de comunicación en un mundo eminentemente oral. Por ejemplo, acceder a los servicios educativos, de salud, justicia, puede resultar estresante sin la ayuda de intérpretes. Diversos sistemas computacionales fueron propuestos para reducir la barrera de comunicación entre personas oyentes y sordas, no obstante, reconocer señas de manera fluida y continua en tiempo real es todavía un grande desafío para la comunidad científica. De acuerdo a la Organización Mundial de Salud, 5% de la población mundial sufre una pérdida de audición incapacitante (466 millones de personas, de ellos aproximadamente 46 millones son niños) (WHO, 2018). En el caso peruano, 1.8% de la población total tiene limitaciones auditiva permanente (CONADIS, 2016). Aquí, presentamos un proyecto de investigación con el objetivo de reconocer de manera automática gestos de la Lengua de Señas Peruana (LSP) utilizada en el área de salud. Siendo un proyecto primigenio, primeramente se recolectará una base de datos de señas utilizando un sensor RGB-D (Red, Green, Blue y Depth), seguidamente se aplicará diferentes técnicas de visión computacional y aprendizaje de máquinas para crear modelos y reconocer tales modelos. Para asegurar que los datos sean reales, se trabajará con personas con necesidades especiales de audición de la Escuela de Educación Especial San Juan de Dios de Ayacucho.

### **PALABRAS CLAVES**

Reconocimiento de lengua de señas, Gestos Aislados de la Lengua de Señas, Fusión de Imágenes, Fusión de Imágenes de color y de profundidad, Gestos en área de salud.

### **1.4. TIPO DE INVESTIGACIÓN**

Prospectivo, longitudinal y analítico.

### **1.5. CRONOGRAMA**

<b>Actividad</b>	<b>1er Trim.</b>	<b>2do Trim.</b>	<b>3er Trim.</b>	<b>4to Trim.</b>
<b>MARCO TEÓRICO</b>				
Introducción	<b>X</b>			

Marco teórico	X			
<b>MATERIAL Y MÉTODOS</b>				
Tipo de investigación		X		
Diseño de la investigación		X		
Población y muestra		X		
Variables e indicadores		X		
Técnicas e instrumentos de recolección de datos		X		
Técnicas y métodos para procesamiento de datos		X		
<b>RESULTADOS</b>				
Análisis y tratamiento de datos			X	
Resultados			X	
Discusión de resultados			X	
<b>CONCLUSIONES Y RECOMENDACIONES</b>				
Conclusiones				X
Recomendaciones				X
<b>PRESENTACIÓN DEL INFORME FINAL</b>				X

#### 1.6. RECURSOS DISPONIBLES

Recursos	Descripción
Humanos	Docentes investigadores
Materiales	Bibliografía, equipos y tecnologías de información
Económicos	De la UNSCH

#### 1.7. PRESUPUESTO

##### 02.0 BIENES

Descripción	Unidad	Cant.	Precio Unitario (S/.)	Precio Total (S/.)
Libros	unidad	10	400.00	4000.00
Artículos científicos	unidad	50	200.00	10000.00
Laptop core i7, 32 Gb, placa gráfica tesla Gb y 4 TB DD	unidad	01	9000.00	9000.00
Camara RGB-D Intel outdoor	unidad	01	1500.00	1500.00
Papel Bond A4 de 80 gr	Millar	05	40.00	200.00
Material de escritorio	Global	01	1000.00	1000.00
<b>Total Bienes (S/.)</b>				<b>25,700.00</b>

##### 03.0 SERVICIOS

Descripción	Unidad	Cant.	Precio Unitario (S/.)	Precio Total (S/.)
Servicio de Internet	Mes	10	150.00	1500.00
Movilidad local	Pasaje	100	10.00	1000.00
Movilidad nacional	Pasaje	16	150.00	2400.00
Viáticos	Día	24	200.00	4800.00
Publicación internacional	Global	03	4000.00	12000.00
Publicación inscripción en conferencia	unidad	03	800.00	2400.00
<b>Total Servicios (S/.)</b>				<b>24,100.00</b>

## **TOTAL DEL PRESUPUESTO (S/): 49,800.00**

### **1.8. FINANCIAMIENTO**

El Financiamiento del Proyecto, a cargo de la Universidad Nacional de San Cristóbal de Huamanga, con el 100% del presupuesto.

## **II. PLAN DE INVESTIGACIÓN**

### **2.1. PROBLEMA**

Diariamente, la mayoría de seres humanos se comunican utilizando alguna lengua oral (LO), mientras que las personas sordo-mudas utilizan una lengua de señas (LS) como sistema de comunicación. La LS utiliza movimientos estandarizados de manos, brazos, postura de cuerpo y expresiones faciales para transmitir mensajes de cualquier naturaleza. Pese a las bondades de la lengua de señas, las personas sordo-mudas enfrentan serios problemas de comunicación en este mundo de comunicación eminentemente oral. Sin la asistencia de intérpretes, por citar algunos ejemplos, sería casi imposible acceder a servicios públicos de educación, atenderse en un puesto de salud, interactuar con representantes de justicia, acceder a un puesto de trabajo, entre otros. Esto es debido a que personas oyentes no saben lengua de señas y las personas sordo-mudas nunca aprendieron a hablar debido a su limitación auditiva, siendo estas últimas, marginadas, ignoradas y aisladas por la sociedad. Diferentes sistemas computacionales de reconocimiento automático de lengua de señas fueron propuestos a fin de mitigar la barrera de comunicación entre personas oyentes y sordas, no obstante, reconocer señas de manera continua y rápida en tiempo real todavía constituye un grande desafío para la comunidad científica (Cooper, Holt, & Bowden, 2011) (Suarez & Murphy, 2012) (Schmidhuber, 2015)(Joudaki et al., 2014).

## **FORMULACIÓN DEL PROBLEMA**

### **PROBLEMA PRINCIPAL**

¿Cómo reconocer gestos aislados de la lengua de señas en tiempo real, área de salud, 2019?

### **PROBLEMAS SECUNDARIOS**

- a. ¿De qué manera reconocer señas simples estáticas en tiempo real?
- b. ¿De qué manera reconocer señas simples dinámicas en tiempo real?

### **2.2. OBJETIVOS**

#### **OBJETIVO GENERAL**

Desarrollar un método para reconocer gestos aislados de la lengua de señas mediante técnicas e instrumentos, técnicas de tratamiento digital de imágenes, visión computacional y aprendizaje de máquinas, con el fin de mejorar el tiempo real de interpretación del gesto, área de salud, 2019

#### **OBJETIVOS ESPECÍFICOS**

- a. Desarrollar un método de fusión de imágenes de color y profundidad para reconocer señas simples estáticas con la finalidad de mejorar el tiempo real de interpretación del gesto.
- b. Desarrollar un método de fusión de imágenes de color y profundidad para reconocer señas simples dinámicas con la finalidad de mejorar el tiempo real de interpretación del gesto.

## **2.3. MARCO TEÓRICO**

### **2.3.1. ANTECEDENTES DE LA INVESTIGACION**

Métodos de reconocimiento automático de lengua de señas toman ventaja de las nuevas cámaras que proveen datos multimodales (color, profundidad y esqueleto del cuerpo) para extraer características de forma de las manos, cuero y expresiones faciales del gesto (Fossati, A., Gall, J., Grabner, H., Ren, X., Konolige, 2013).

Recientemente, (Yauri Vidalón & De Martino, 2016) presentaron un trabajo para reconocer gestos aislados de la lengua de señas brasileña utilizando DTW y descriptores basados solamente en imágenes de profundidad. Trabajando con 107 señas del área médica consiguieron una tasa de reconocimiento de 98.69% en una base de datos dependiente del señalizador.

### **2.3.2. GESTOS AISLADOS DE LA LENGUA DE SEÑAS**

Los gestos son una forma de comunicación visual no verbal, realizado por movimientos de manos, brazos, cabeza, expresiones faciales y posturas del cuerpo. Estos poseen gran expresividad, pudiendo revelar pensamientos, intenciones, entre otros (Abdelkader, Abd-Almageed, Srivastava, & Chellappa, 2011).

Los gestos pueden ser expresados en forma aislada y continua (Cooper, Holt, & Bowden, 2011). Un gesto simple se realiza de forma aislada, por tanto sus fronteras de inicio y fin son fácilmente identificables. Un gesto aislado no es afectado por el signo precedente ni por el signo sucesivo. En el otro lado, en la expresión continua, los gestos son expresados uno después del otro, en secuencia, sin una frontera clara entre las mismas.

En ausencia del habla, los gestos pueden constituir un medio de comunicación, como sucede con la lengua de señas (Valli, 2000) (Sandler & Lillo-Martin, 2006). Es más, los gestos en una lengua de señas son bien estructurados y estandarizados, sujetos a fuertes reglas de contexto y gramática, por lo que son susceptibles de ser reconocidos mediante técnicas computacionales (Kuznetsova, Leal-Taixe, & Rosenhahn, 2013) (Fossati, A., Gall, J., Grabner, H., Ren, X., Konolige, 2013) (Yauri Vidalón & De Martino, 2016).

### **SEÑA AISLADA ESTÁTICA**

Un gesto o seña de la lengua de señas puede ser descrita en función a parámetros de configuración, localización y trayectoria de movimiento de la mano (o manos), junto con otras características no manuales (Valli, 2000). Una seña aislada estática consta de una postura fija, sin ningún cambio tanto en las características manuales y no manuales del gesto

### **SEÑA AISLADA DINÁMICA**

Una seña aislada dinámica es la opuesta a una seña estática. La realización del gesto implica cambios en la configuración, localización y trayectoria de movimiento de las manos y características no manuales.

### **2.3.3. TIEMPO REAL**

Una respuesta en tiempo real garantiza una respuesta dentro de las restricciones de tiempo especificadas, denominado tiempo límite.

De acuerdo con (Gnome 2012), el tiempo de respuesta aceptable de sistema interactivo varía con las tareas. Así, para eventos que aguarden respuestas de los usuarios, el tiempo límite es de 10 segundos.

### **CUADROS POR SEGUNDO**

Número de cuadros o imágenes que serán procesados para cumplir el criterio de tiempo real en nuestro sistema.

### **COMPLEJIDAD COMPUTACIONAL**

A fin de realizar nuestra comparación de manera independiente del sistema computacional, se calculará la complejidad computacional de nuestra propuesta, siendo el número de pasos requeridos.

## **2.4. HIPÓTESIS**

### **HIPÓTESIS GENERAL**

La fusión de imágenes para reconocer gestos aislados de la lengua de señas mejora el tiempo real de interpretación del gesto.

### **HIPÓTESIS ESPECÍFICAS**

- a. La fusión de imágenes de color y profundidad para reconocer señas simples estáticas mejora el tiempo real de interpretación del gesto.
- b. La fusión de imágenes de color y profundidad para reconocer señas simples dinámicas mejora el tiempo real de interpretación del gesto.

## **2.5 VARIABLES E INDICADORES**

### **Variable Independiente**

X: Gestos aislados de la lengua de señas.

### **Indicadores**

X1: Seña aislada estática.

X2: Seña aislada dinámica.

### **Variable Dependiente**

Y: Tiempo real.

### **Indicadores**

Y1: Cuadros por segundo.

Y2: Complejidad computacional.

## **2.6 DISEÑO METODOLÓGICO**

### **TIPO DE INVESTIGACIÓN**

Prospectivo, longitudinal y analítico.

### **NIVEL DE INVESTIGACIÓN**

Explicativo.

### **MÉTODO**

- a) Inductivo.
- b) Deductivo.
- c) Análisis.
- d) Síntesis.
- e) Matemático.

f) Estadístico.

## **DISEÑO**

Experimental.

## **POBLACIÓN**

La lengua de señas peruana de personas sordas y mudas.

## **MUESTRA**

La lengua de señas peruana de personas sordas y mudas que buscan servicios de salud.

## **TÉCNICAS**

- a) Entrevista
- b) Grabación de video.

## **INSTRUMENTOS**

- a) Guía de entrevista.
- b) Equipos de grabación.

## **2.7 REFERENCIAS BIBLIOGRÁFICAS**

1. Abdelkader, M. F., Abd-Almageed, W., Srivastava, A., & Chellappa, R. (2011). Silhouette-based gesture and action recognition via modeling trajectories on Riemannian shape manifolds. *Computer Vision and Image Understanding*, 115(3), 439–455. <https://doi.org/10.1016/j.cviu.2010.10.006>
2. Bishop, C. M. (2006). *Pattern Recognition and Machine Learning* (Information Science and Statistics). Berlin, Heidelberg: Springer-Verlag.
3. CONADIS. (03 de 12 de 2016). Situación de las Personas con Discapacidad Auditiva. Obtenido de Estadísticas en Discapacidad: <https://www.conadisperu.gob.pe/estadisticas-en-discapacidad>
4. Cooper, H., Holt, B., & Bowden, R. (2011). Sign Language Recognition. In T. B. Moeslund, A. Hilton, V. Krüger, & L. Sigal (Eds.), *Visual Analysis of Humans* (pp. 539–562). London: Springer London. <https://doi.org/10.1007/978-0-85729-997-0>
5. Fossati, A., Gall, J., Grabner, H., Ren, X., Konolige, K. (Ed.). (2013). *Consumer Depth Cameras for Computer Vision - Research Topics and Applications*. Springer Publishing Company.
6. Gonzalez, R. C., & Woods, R. E. (2006). *Digital Image Processing* (3rd Edition). Upper Saddle River, NJ, USA: Prentice-Hall, Inc.
7. Gnome. (03 de 12 de 2018). Gnome Human Interface Guidelines 2.2.2. Acceptable Response Times. Obtenido de <https://developer.gnome.org/hig-book/3.2/feedback-response-times.html.en>
8. Joudaki, S., Mohamad, D. bin, Saba, T., Rehman, A., Al-Rodhaan, M., & Al-Dhelaan, A. (2014). Vision-Based Sign Language Classification: A Directional Review. *IETE Technical Review*, 31(5), 383–391. <https://doi.org/10.1080/02564602.2014.961576>
9. Kuznetsova, A., Leal-Taixe, L., & Rosenhahn, B. (2013). Real-Time Sign Language Recognition Using a Consumer Depth Camera. In 2013 IEEE International Conference on Computer Vision Workshops (pp. 83–90). IEEE. <https://doi.org/10.1109/ICCVW.2013.18>
10. Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. The MIT Press.
11. Sandler, W., & Lillo-Martin, D. (2006). *Sign Language and Linguistic*



- Universals. Cambridge University Press.
12. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
  13. Suarez, J., & Murphy, R. R. (2012). Hand gesture recognition with depth images: A review. In 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication (pp. 411–417). IEEE. <https://doi.org/10.1109/ROMAN.2012.6343787>
  14. Valli, C. (2000). *Linguistics of American Sign Language: An Introduction*. Gallaudet University Press. Retrieved from <http://books.google.com/books?id=mfS3GITLAUMC&pgis=1>
  15. Yauri Vidalón, J. E., & De Martino, J. M. (2016). Brazilian Sign Language Recognition Using Kinect (pp. 391–402). Springer International Publishing. [https://doi.org/10.1007/978-3-319-48881-3\\_27](https://doi.org/10.1007/978-3-319-48881-3_27)
  16. Who. (03 de 12 de 2018). World Health Organization - Deafness and hearing loss. Obtenido de <http://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>

### III. ANEXOS

#### 3.1. MATRIZ DE CONSISTENCIA

**TITULO:** Método para Reconocimiento de Gestos Aislados de la Lengua de Señas en Tiempo Real, Área de Salud, 2019.

**RESPONSABLE:** Mg. Ing. Manuel Avelino Lagos Barzola - **MIEMBROS:** MSc. Ing. Efraín Elías Porras Flores; MSc. Ing. José Elías Yauri

Vidalón. **COLABORADOR** : Ing. Lisber Arana Hinostroza

PROBLEMA	OBJETIVOS	MARCO TEÓRICO	HIPÓTESIS	VARIABLES E INDICADORES	METODOLOGÍA
<p><b>PROBLEMA PRINCIPAL</b> ¿Cómo reconocer gestos aislados de la lengua de señas en tiempo real, área de salud, 2019?.</p> <p><b>PROBLEMAS SECUNDARIOS</b> a. ¿De qué manera reconocer señas simples estáticas en tiempo real?.</p> <p>b. ¿De qué manera reconocer señas simples dinámicas en tiempo real?.</p>	<p><b>OBJETIVO GENERAL</b> Desarrollar un método para reconocer gestos aislados de la lengua de señas mediante técnicas e instrumentos, técnicas de tratamiento digital de imágenes, visión computacional y aprendizaje de máquinas, con el fin de mejorar el tiempo real de interpretación del gesto, área de salud, 2019</p> <p><b>OBJETIVOS ESPECÍFICOS</b> a. Desarrollar un método de fusión de</p>	<p><b>ANTECEDENTES</b> Métodos de reconocimiento automático de lengua de señas toman ventaja de las nuevas cámaras que proveen datos multimodales para extraer características de los gesto (Fossati, A., Gall, J., Grabner, H., Ren, X., Konolige, 2013).</p> <p>Recientemente, (Yauri Vidalón &amp; De Martino, 2016) presentaron un trabajo para reconocer gestos aislados de la lengua de señas brasileña utilizando DTW y descriptores basados solamente en imágenes de profundidad. Trabajando con 107 señas del área médica</p>	<p><b>HIPÓTESIS GENERAL</b> La fusión de imágenes para reconocer gestos aislados de la lengua de señas mejora el tiempo real de interpretación del gesto.</p> <p><b>HIPÓTESIS ESPECÍFICAS</b> a. La fusión de imágenes de color y profundidad para reconocer señas simples estáticas mejora el tiempo real de interpretación del gesto. b. La fusión de</p>	<p><b>VARIABLE INDEPENDIENTE</b> X: Gestos aislados de la lengua de señas.</p> <p><b>INDICADORES</b> X1: Señal aislada estática. X2: Señal aislada dinámica.</p> <p><b>VARIABLE DEPENDIENTE</b> Y: Tiempo real.</p> <p><b>INDICADORES</b> Y1: Cuadros por segundo. Y2: Complejidad computacional.</p>	<p><b>TIPO DE INVESTIGACIÓN</b> Prospectivo, longitudinal y analítico.</p> <p><b>NIVEL DE INVESTIGACIÓN</b> Aplicativo.</p> <p><b>MÉTODO</b> a) Inductivo. b) Deductivo. c) Análisis. d) Síntesis. e) Matemático. f) Estadístico.</p> <p><b>DISEÑO</b> Experimental.</p> <p><b>POBLACIÓN</b> La lengua de señas peruana de personas sordas y</p>

	<p>imágenes de color y profundidad para reconocer señas simples estáticas con la finalidad de mejorar el tiempo real de interpretación del gesto.</p> <p>b. Desarrollar un método de fusión de imágenes de color y profundidad para reconocer señas simples dinámicas con la finalidad de mejorar el tiempo real de interpretación del gesto.</p>	<p>consiguieron una tasa de reconocimiento de 98.69% en una base de datos dependiente del señalizador.</p> <p><b>GESTOS AISLADOS DE LA LENGUA DE SEÑAS</b> Un gesto aislado se realiza de forma separa e independiente, por tanto sus fronteras de inicio y fin son fácilmente identificables. Por tanto, un gesto aislado no es afectado por el signo precedente ni por el signo sucesivo(Cooper, Holt, &amp; Bowden, 2011).</p> <p><b>TIEMPO REAL</b> De acuerdo con (Gnome 2012), el tiempo de respuesta aceptable de sistema interactivo varía con las tareas. Así, para eventos que aguarden respuestas de los usuarios, el tiempo límite es de 10 segundos.</p>	<p>imágenes de color y profundidad para reconocer señas simples dinámicas mejora el tiempo real de interpretación del gesto.</p>		<p>mudas.</p> <p><b>MUESTRA</b> La lengua de señas peruana de personas sordas y mudas que buscan servicios de salud.</p> <p><b>TÉCNICAS</b> a) Entrevista b) Grabación de video.</p> <p><b>INSTRUMENTOS</b> a) Guia de entrevista. b) Equipos de grabación.</p>
--	---	---	--	--	---